# Novel β-Hairpin Peptidomimetic Inhibitors of Protein-RNA Interactions in BIV/HIV-1 and in Bacterial RNase P

**Dissertation**

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

**Krystyna Patora-Komisarska**

aus Polen

Promotionskomitee

Prof. Dr. John A. Robinson (Vorsitz)

Prof. Jay S. Siegel

Zürich 2008

## Summary

The aim of the research described in this thesis is the development of novel antiviral and antibacterial β-hairpin peptidomimetics. The focus has been on the inhibition of the critical protein-RNA interactions between the viral regulatory proteins Tat and Rev and their target RNAs TAR and RRE, respectively, and the RNA and protein subunits of bacterial RNase P. RNA and its interactions with proteins has grown greatly in importance during the last decade. The increasing resistance against antiviral and antibacterial drugs puts these novel targets into a favourable light, as it is generally thought that pathogens will find it difficult to mutate RNAs that fulfill essential functions. Efforts have been made to mimic the native RNA binding domains of the above-mentioned proteins using cyclic peptides containing a heterochiral diproline template. These peptides are an interesting class of protein mimetics due to the strong β-hairpin inducing properties of the D-Pro-L-Pro template. They can be made efficiently using commercially available building blocks and standard solid-phase and solution synthesis methods.

The interaction between the viral protein Tat and its cognate transactivator response element (TAR) RNA is critical for the life cycle of the virus. The bovine immunodeficiency virus (BIV) Tat arginine-rich RNA binding domain is found to adopt a β-hairpin conformation upon binding to TAR RNA. Previous studies on BIV Tat peptidomimetics allowed the identification of one highly potent BIV Tat-TAR inhibitor, called BIV2. Its stable β-hairpin conformation and high affinity to BIV TAR RNA enabled determination of the structure of its complex with BIV TAR by NMR techniques. In this work, extensive structure-activity relationship (SAR) studies on these 12-mer loop BIV Tat peptidomimetics were carried out, which yielded further highly potent and selective ligands. Furthermore, the BIV Tat-TAR interaction was employed as a model from which important information about RNA recognition was derived and exported to human immunodeficiency virus type-1 (HIV-1) Tat-TAR, facilitating the design of novel β-hairpin peptidomimetic inhibitors of HIV Tat. From assaying an extensive library of BIV Tat peptidomimetics, some potent and selective ligands for HIV-1 TAR RNA have been identified, the best of which **L2-15,** exhibits a dissociation constant of $K_d$~75 nM and an $IC_{50}$~100 nM for inhibition of ADP-1

binding to TAR. Analogous to the BIV system, the HIV-1 Tat protein binds to the bulge region of the TAR RNA stem-loop structure. However, in order to function, a ternary complex between Tat-TAR and cyclin T1 must be formed, in which cyclin T1 interacts with Tat and the tip of the RNA TAR loop. Guided by computational modelling and SAR studies, novel 16-mer loop HIV-1 Tat peptidomimetics were developed, which exhibit a stable β-hairpin conformation in aqueous solution, bind with high affinities to TAR RNA and are able to inhibit the Tat-TAR interaction. The best peptide in this class, **L5-11**, exhibits a dissociation constant of $K_d$~25 nM in the presence of a large excess of tRNA, and $K_d$~0.5 nM in the absence of the competitor RNA. Since the structure of the HIV-1 Tat-TAR complex is not yet available, extensive SAR studies on the 16-mer loop mimetics have been performed, in order to deepen understanding of the recognition principles in this protein-RNA interaction. Further studies on HIV Tat peptidomimetics established their ability to internalise through the cell membrane and their stability in biological settings.

A second critical step for viral replication is binding of Rev protein to Rev response element (RRE) RNA, which facilities successful transport of full length or partially spliced mRNA from the nucleus into the cytoplasm. The RNA-binding domain of Rev protein was found previously to form an α-helical structure upon binding to its target RNA. In the second part of this thesis, a novel approach is presented to mimic this α-helical structure using a β-hairpin peptidomimetic, which provides a robust scaffold upon which the groups critical for RRE recognition are displayed. First, assaying the already available related β-hairpin peptidomimetics based on BIV Tat-derived peptide, and then designing a new series of peptides based on the sequence of the RNA-binding domain of Rev, gave a set of potent RRE ligands, with $K_d$'s comparable to that of the native Rev peptide. The hairpin conformation of the peptides was stabilized by introducing a disulfide cross-link, resulting in a very active Rev peptidomimetic, **R2-14**, with a stable β-hairpin structure in aqueous solution.

A similar approach is also described for the design of novel peptidic inhibitors of *E. coli* RNase P, a ribonucleoprotein, which facilitates the maturation of precursor tRNA in all three domains of life. The extensive biochemical studies on two related bacterial strains, *E. coli* and *B. subtilis*, revealed that α-helical, highly conserved and arginine-rich regions of the protein subunit of RNase P interact with the catalytically active RNA subunit. By assaying peptides from the Rev library, exhibiting sequence

similarities to the RNA-binding domain of RNase P protein, two potent antimicrobial agents acting selectively on *E. coli* were discovered. Combining the information from these lead compounds and preparing a computational model of a β-hairpin superimposed on the helical epitope of interest, led to the discovery of highly active inhibitors of RNase P in *E. coli*, the best of which, **P2-05**, possessed a MIC value of 0.06 μg/ml. This class of peptides has proved to have favourable *in vitro* pharmacological properties including low haemolytic activity and high stability in plasma, which makes them interesting as prospective antimicrobial agents.

All three above-mentioned approaches to inhibitors of protein-RNA interactions resulted in the discovery of very potent and selective antiviral and antibacterial agents that are very attractive starting points for drug development. In addition to their pharmaceutical potential, the extensive SAR studies provide important information about the chemistry and principles governing RNA recognition by these structurally and conformationally constrained peptidomimetics, which might aid the optimisation of other known RNA ligands, and in the future, the development of agents targeting other functional RNAs.

## Table of contents

# 1. INTRODUCTION

## 1.1. RNA - a target for drug discovery

The growing awareness of the remarkable variety of biochemical roles of RNA in all living organisms is leading to an increasing appreciation that cellular and viral RNAs provide inviting targets for the treatment of both infectious and chronic diseases. Since the emergence of drug resistance is an ever-increasing therapeutic problem in traditional antibacterial and antiviral therapies, exploiting a new class of molecules that are chemically distinct from protein-based targets, namely RNA, offers a promising strategy to overcome this problem.

RNA has for a long time been a validated drug target for antibacterial treatment. Antibiotics that have been in clinical use for decades (e.g. erythromycin) bind ribosomal RNA.[1] Furthermore, in the last decade, RNA interference (RNAi), a technique that utilizes small interfering RNAs (siRNAs) to specifically knock-down the expression of targeted genes at the post-transcriptional level, has gained a lot of interest. This methodology not only has broad application in basic biomedical research, but is also being used in the development of therapeutic agents.[2] A milestone in this field was the approval of a first antisense drug in 1998, Vitravene, to treat a condition called cytomegalovirus (CMV) retinitis in people with AIDS.

Most functions of RNA require an interaction with RNA-binding proteins. Therefore protein-RNA complexes are not only important for understanding RNA recognition, but also protein-binding sites in RNA are potential targets for drug discovery. Prominent examples are interactions critical for virus replication in human immunodeficiency virus (HIV) between the regulatory proteins Tat and Rev and their cognate RNAs, TAR and RRE, respectively.

### 1.1.1. RNA-binding proteins

Specific RNA-protein interactions play crucial roles in gene regulation through transcriptional control, RNA processing, RNA transport and translation. During transcription, RNA structures mediate several important regulatory mechanisms, such

as transcription termination by the eubacterial mRNA-binding Rho factor,[3] the tRNA-dependent transcription anti-termination in eubacteria,[4] and the *trans*-activation of viral genome transcription in HIV and related retroviruses.[5] Primary RNA transcripts undergo a variety of different processing and localization events, depending on the type of RNA (mRNA, rRNA, tRNA, etc.) and organism. The three main post-transcriptional modifications of mRNA in eukaryotic cells are: 5'-capping, 3'-polyadenylation, and RNA splicing.[6] In the nucleus, pre-mRNA is associated with a variety of proteins in complexes known as heterogeneous ribonucleoprotein particles (hnRNPs).[7] Spliceosomes, to mention just one example of such pre-mRNA-protein complexes, are essential to remove internal non-coding regions of pre-mRNA (introns) and to join the remaining segments (exons) into mRNA before translation.[8] After maturation, mRNA has to be transported from the nucleus to the cytoplasm, where translation can take place. This process is facilitated by nuclear pores, large protein complexes that cross the nuclear envelope. In HIV-1, for example, the genomic RNA is exported with the help of the Rev protein, which has a specific binding site for the RRE motif present in unspliced and partially-spliced HIV-1 RNA and an NES motif (nuclear export signal), which interacts with exportin-1. Another essential RNA-protein complex common for all the life kingdoms is RNase P. This ribonuclease catalyses endonucleotic cleavage of the 5'-leader sequence of precursor tRNAs, producing a mature 5'-phosphate at the end of the tRNA.

The activity of the family of RNA-binding proteins is mediated by the specific recognition of sequences and structural elements within the mRNA transcripts. RNA is normally single stranded but often forms secondary structures, such as hairpins (stem-loops), bulges and internal loops through pairing of complementary stretches within a strand. These secondary structural elements are often used as binding sites for proteins. Characterization of these RNA binding proteins resulted in the identification of several RNA-binding motifs, including RNA-recognition motifs (RRMs), K-homology domains (KH), Zinc finger (ZnF), double-stranded RNA binding domains (dsRBD), and arginine–rich motifs.[9,10]

The RNA-recognition motif (RRM, also referred to as RNA binding domain (RBD) or RNP-motif) is by far the most common and best characterised of the RNA-binding molecules. It is usually composed of 80-90 amino acids that form a four stranded antiparallel β-sheet with two helices packed against it, and follows a general

pattern of β–α–β–β–α–β. In the approximately 20 known structures of RRM-RNA complexes, RNA recognition usually occurs on the surface of the first β-sheet. Binding is mediated in most cases by three conserved residues: an Arg or Lys residue that forms a salt bridge to the phosphodiester backbone and two aromatic residues that make stacking interactions with the nucleobases. These three amino acids reside in the two highly conserved motifs, RNP motif-1 (RNP1) and RNP2, and define these motifs at the sequence level and are located in the two central β-strands.[11,12]

K-homology (KH) protein is a domain that binds both single-stranded DNA and RNA and is ubiquitous in eucaryotes, eubacteria and archea. The domain is composed of 60-70 amino acids with a functionally important signature sequence of (I/L/V)IGXXGXX(I/L/V) near the centre of the domain. The KH-domain contains three antiparallel β-sheets, which are packed against three α-helices on one face of the β-sheet (β–α–α–β–β–α or α–β–β–α–α–β) The loops between β2/β3 or α2/β2 as well as a GXXG motif are important for recognizing about four nucleotides in single stranded RNA (ssRNA).[13,14]

The double-stranded RNA binding domain, dsRBD, is another small domain composed of 60-70 amino acids, which folds into a structure where two α-helices lie on the face of a three-stranded antiparallel β-sheet (α–β–β–β–α). Such domains are found in both bacteria and eucaryotes. Helix α1, the N-terminal portion of α2-helix and the loop between β1 and β2 are important for the shape-specific recognition of the minor-major-minor groove of dsRNA through contacts to the sugar-phosphate backbone.[15]

Zinc-fingers (ZnF) are one of the most versatile binding motifs, first known as DNA-binding domains, but now also known to have additional activities such as the recognition of RNA and other proteins. They are characterized by a Cys- and His-rich domain that chelates a zinc ion. Different classes of zinc-finger proteins contain different combinations of metal binding amino-acids: CCHH (Cys-Cys-His-His), CCCH or CCHC. Structural studies of ZnF complexes reveal considerable diversity in terms of protein partners, binding modes and affinities, and highlight the often underestimated versatility of the ZnF structure and function (Figure 1).[16,17]

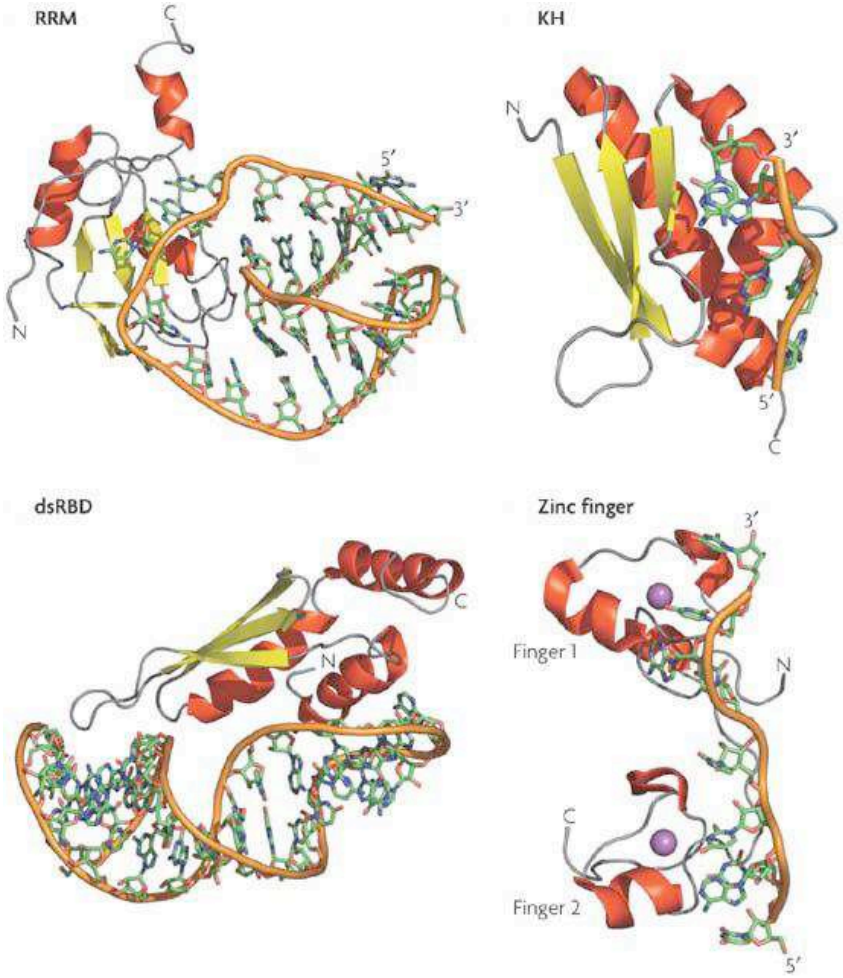
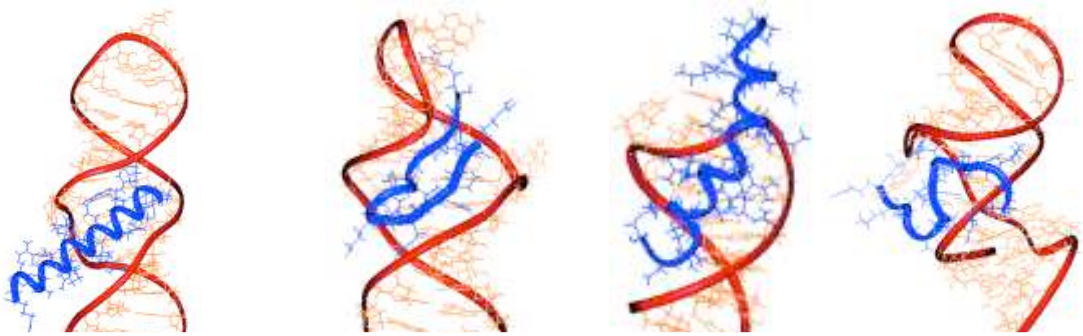



**Figure 1.** RNA-binding proteins: **RRM**-Structure of the N-terminal RNA-recognising motif of human U1A bound to RNA;[12] **KH**-domain of Nova-2 bound to 5'-AUCAC-3';[18] **dsRNA**-binding domain from yeast Rnt1 bound to an RNA helix;[19] two **ZnF** of TlS11d bound to an AU-rich RNA element.[20]

In contrast to the prefolded RNA-binding domains described above, short arginine-rich domains, called also arginine-rich motifs (ARM), adopt stable conformations only upon binding to their target RNAs. Proteins containing ARMs function in transcription, translation, RNA trafficking and packaging and are present in anti-termination bacteriophage proteins (bacteriophage λ N protein),[21,22] viral regulatory proteins (Tat, Rev, Rex), ribosomal proteins[23] and in naturally occurring antimicrobial peptides.[24]



HIV Rev-RRE RNA α-helix | BIV Tat-TAR RNA β-hairpin | λ N-box B RNA bent α-helix | Rex-RNA aptamer extended conformation

**Figure 2.** Arginine-rich binding domains, populate a variety of conformations. NMR structures of peptide-RNA complexes: **HIV Rev-RRE** interaction,[25] bovine immunodeficiency virus (BIV) **Tat-TAR** interaction,[26] bacteriophage **λ N peptide-Box B RNA**,[27] human T-cell leukemia virus (HTLV-1), **Rex-RNA** aptamer.[28]

The ARMs are usually 20-30 amino acids long and, while unfolded or weakly structured in the absence of RNA, have been shown to bind RNA in α-helical, β-hairpin and other conformations (Figure 2). Although it is not surprising that highly positively charged peptides derived from this class of RNA-binding proteins can bind RNA with a high affinity, it is remarkable that they bind with relatively high specificity. This, apparently, is possible because RNAs form distinct binding pockets that help to induce peptide conformations and make specific interactions including salt bridges, charge-stabilized and neutral hydrogen bonding, weakly polar and π–stacking interactions and precise apposition of van der Waals interactions. Adaptive binding of peptide secondary structures is associated with widening of the RNA major groove to accommodate the ligand. The molecular mechanisms of the major-groove widening are different for each complex but share the use of non-canonical RNA elements.

### 1.1.2. RNA-binding domains as attractive targets for drug design

The ability of RNA molecules to fold into complex three-dimensional structures, and the importance of RNAs to many biological processes, have generated great interest in designing sequence-specific RNA binding molecules. One advantage of targeting RNA over traditional protein targets might be slower development of drug resistance. The RNA functional domains, potential drug targets, are more highly conserved, and it is expected that pathogens will find it more difficult to mutate their RNA. The complexity of RNA structure, including double-stranded duplexes, hairpins,

internal loops, bulged bases, and pseudo-knotted structures, makes it difficult to design ligands for sequence-specific RNA recognition. However, de novo design can be very useful for testing our understanding of the basic principles of nucleic acid recognition, as well as developing new types of drugs to treat both infectious and chronic diseases. So far, RNA is a validated drug target for antibacterial treatment. Antibiotics, such as aminoglycosides (streptomycin, tombramycin, neomycin), polyketides (tetracycline) and macrolides (erythromycin), were all discovered in the ca. 1950's and target bacterial ribosomal subunits, thus inhibiting bacterial protein synthesis.[1,29] Furthermore, RNA is the genetic material of pathogenic viruses such as HIV or hepatitis C virus (HCV), so it provides numerous opportunities for the discovery of new drugs to treat these diseases. Finally, the complex functions of RNA in the control of gene expression in humans opens many new opportunities to target specific RNAs for the treatment of numerous chronic health conditions. To date, many of the RNA targets have been approached by modifying existing ribosomal antibiotics. However, new technologies such as antisense oligonucleotide agents, siRNA molecules to induce gene silencing by RNA interference, have been extensively exploited. However, drawbacks of antisense technology are the poor pharmacological profile, unspecific protein binding, inefficient metabolic stability and poor cellular uptake. In addition, many RNA sequences are not accessible to antisense agents because they are highly structured or bound to cellular proteins.[30,31]

Among many RNA targets used for drug design, in this work the focus is on RNA interacting with arginine-rich domains, including viral BIV/HIV-1 Tat-TAR RNA and HIV Rev-RRE RNA interactions, and bacterial RNase P holoenzyme. These targets provide good scaffolds for inhibitor design, since the proteins concerned are short, conformationally diverse, but well studied from a structural viewpoint and they bind to RNA with high affinities and specificities. In the literature, mostly two strategies have been taken towards the design of RNA-protein interaction inhibitors: **i)** by modifications of known RNA-binding natural products, such as aminoglycosides, or **ii)** imitating the structure of a native binding domain by short, usually constrained peptides. Aminoglycosides seem to be a very attractive starting point, since several structures of aminoglycoside-RNA complexes have been solved and the interactions with RNA have been well defined, particularly for 16S and 30S ribosomal bacterial RNA.[1,32] Although, there are examples of aminoglycosides inhibiting RNase P, or viral Tat-TAR and Rev-RRE interactions, the structures of the RNA-ligand complexes have

been only poorly studied, which makes further optimisation difficult.[33] Furthermore, the poor oral absorption of highly positively charged aminoglycosides, the growing resistance to them, their toxicity and their chemical complexity are drawbacks in the use and development of aminoglycosides and their derivatives. Much effort has been made to develop other RNA-binding small molecules with lower structural complexity than aminoglycosides, and with better pharmacological properties, including oxazolidines, the only fully synthetic antibiotics that target rRNA of the bacterial ribosome, and phenyl-furan derivatives.[29]

Another strategy, the design of relatively small peptidomimetics that can mimic the structure and/or function of the active sites within proteins, is gaining popularity at present. Peptidomimetics are useful for detailed studies of protein folding, structure and function. The main potential application of peptidomimetics is to block protein-protein or protein-nucleic acid interactions involved in undesired biological processes. Being relatively small, peptidomimetics often solve some of the problems associated with the use of proteins as drugs, such as antigenicity. However, low bioavailability and rapid enzymatic degradation remain major problems. The properties of the parent binding-domain must be retained when designing peptidomimetics, including the bioactive conformation of the desired active site, and a certain degree of conformational freedom to allow an induced fit. Linear peptides are not optimal candidates to mimic proteins, because they equilibrate between multiple conformations and their adaptation to the bioactive conformation incurs an entropic cost. Introduction of conformational constraints into the peptide is, therefore, often needed to generate a successful peptidomimetic. Being relatively small and conformationally constrained, cyclic peptides are excellent candidates to serve as potential inhibitors of protein-RNA interactions.[34,35]

### 1.1.3. Specificity and stereospecificity of RNA-ligand binding

A major challenge in the development of RNA-targeted therapeutics is to design molecules binding selectively to a desired RNA target and not to a multitude of other cellular RNAs.[36] A deeper understanding of RNA folding, structure and recognition of specific targets is necessary to enable the successful design of target-selective molecules. Aminoglycosides are among the most commonly used broad-spectrum antibiotics in the clinic. Since the initial discovery of streptomycin in 1944, a series of

aminoglycoside antibiotics, including kanamycin, neomycin B, gentamicin, and tobramycin have been successfully introduced as treatments for serious systemic infections, and have been used for a variety of topical applications.[1] Due to their broad use in clinical applications the aminoglycosides are the most studied RNA-ligands. Aminoglycoside antibiotics are highly selective in their preferential binding of RNA over DNA, but exhibit much less discrimination between different RNA molecules. Aminoglycosides bind to a wide range of unrelated RNAs, including simple duplex RNA,[37] 16S and 18S rRNAs,[38,39] mRNA transcripts,[40] tRNA,[41] and a variety of catalytic RNAs.[42] The poor RNA-target selectivity of the aminoglycosides has been attributed to two major factors, **i)** their highly charged nature, which is responsible for their electrostatically driven binding to the major groove of duplex RNA, and **ii)** their conformational flexibility. The mutual electrostatic attraction between multi-positively charged groups of aminoglycosides and the negatively charged RNA backbone may also lead to base pairing disruption. They may also bind less extensively to human 18S rRNA causing severe toxicity.[1,43,44]

A comparison of the selected, clinically used aminoglycosides that can inhibit the interactions of viral RNAs TAR and RRE with their native ligands is summarised in Table 1. Aminoglycoside antibiotics bind their therapeutic target-bacterial rRNA in the sub-nanomolar/low micromolar range, but due to their low specificity they also interfere with human rRNA, causing toxicity.[45,46] As shown in Table 1 they also bind to other RNA domains, including HIV TAR and RRE RNAs. It is reasonable to suggest that the inherent flexibility of RNA and the aminoglycosides may allow for a myriad of different binding solutions, which lead to the same functional consequences and as a result to low selectivity.

| **Aminoglycoside** | **Clinical Use** | **Target** | **Inhibitory activity** | |
|---|---|---|---|---|
| | | | **HIV Tat-TAR** | **HIV Rev-RRE** |
| Neomycin **(1)** | Excellent against Gram-positive; partially against Gram-negative, skin infections | Binds to the 16S subunit of the bacterial ribosome | $IC_{50} \approx 1\ \mu M$ | $IC_{50} \approx 0.1\ \mu M$ |
| Streptomycin **(2)** | Gram-negative bacteria, tuberculosis | Binds to the 16S subunit of the bacterial ribosome | $IC_{50} \approx 10\ \mu M$ | $IC_{50} \geq 100\ \mu M$ |
| Gentamicin **(3)** | Gram-negative bacteria | Binds the 30S subunit of the bacterial ribosome | $IC_{50} \approx 45\ \mu M$ | $IC_{50} \approx 10\ \mu M$ |

**Table 1.** The comparison of aminoglycoside antibiotics in clinical use interfering with viral RNAs. Also shown are inhibitory values ($IC_{50}$) reported for Tat-TAR RNA from Krebs,[47] and for RRE-Rev from Zapp.[48]

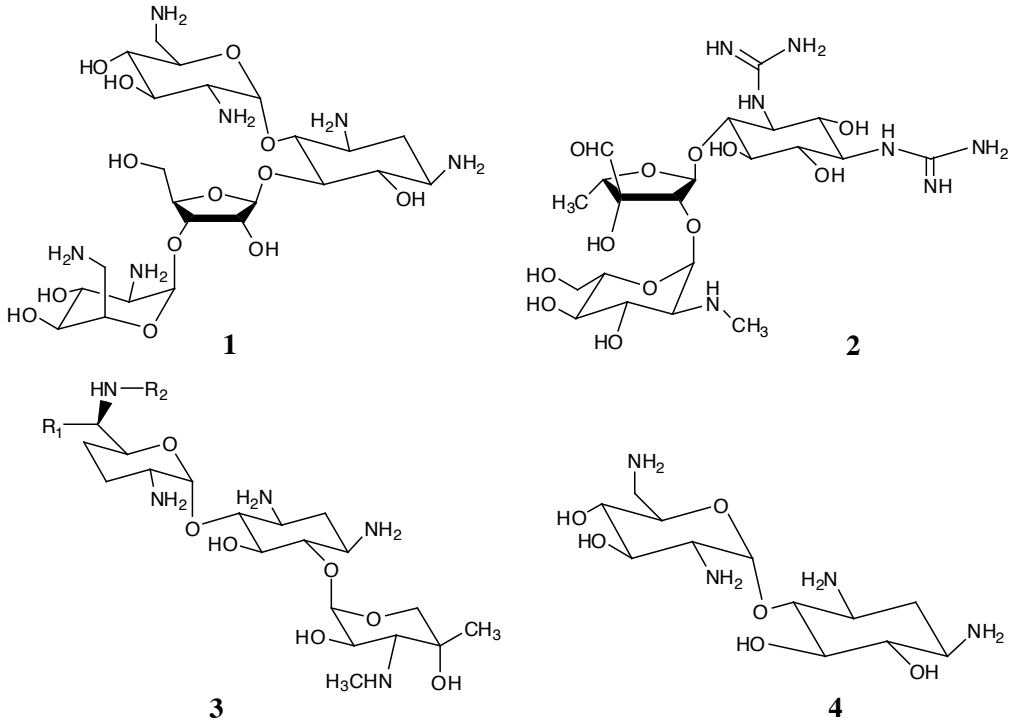



**Figure 3.** Structures of the aminoglycosides that bind to A-site of bacterial rRNA as well as to other unrelated RNAs: neomycin (1), streptomycin (2), gentamicin (3), neamine (4).

The cited results indicate that aminoglycoside-based ligands are capable of high affinity binding to RNA and therefore are interesting compounds for drug discovery, but achieving high site specificity remains a challenging objective.

The same conclusions were drawn from experiments performed with L- and D-neamine **(4,** Figure 3),[49] HIV-1 D/L-Tat peptides[50,51] and D/L-Rev peptides,[52] which showed very low, if any, stereospecificity in binding to the target RNA (Table 2). The enantiomeric forms of these ligands show similar binding modes, including binding to the bulge region of RNA and widening of the major groove.

| Ligand | Target | $K_d$ [nM] |
|---|---|---|
| D-neamine | A-site bacterial rRNA | 501.7±14.2 |
| L-neamine | A-site bacterial rRNA | 1039.5±29.1 |
| L-Tat peptide | TAR RNA | 130±80 |
| D-Tat peptide | TAR RNA | 220±50 |
| L-Rev peptide | RRE RNA | 14.4±0.6 |
| D-Rev peptide | RRE RNA | 14.6±0.3 |

**Table 2**. A comparison of binding affinities of enantiomeric forms of neamine, HIV-1 Tat and Rev peptides to their target RNAs.[49,50,52]

These findings strongly support the view that both partners in the ligand-RNA interaction are mouldable. Thus, rather than being modelled as a rigid key-and-lock, or even as a hand-in-glove fitting, it is suggested that both RNA and ligand undergo conformational changes by a mutual induced-fit process.

In an attempt to circumvent this conformational flexibility and the resulting RNA target promiscuity, conformationally constrained molecules seem to be one of the potential solutions. Freezing the ligand conformation may, under ideal circumstances, have advantages including, **i)** in increased affinity to the desired target due to limited entropy losses upon binding and, **ii)** in increased selectivity by locking the ligand skeleton in an unfavourable orientation for binding to competing targets. This approach has been taken already in the design of conformationally constrained aminoglycosides[53] and peptidomimetics.[54] These studies show that designing a constrained compound that binds with the same affinity as the parent natural protein is not a trivial task, however, by developing structure-activity-relationships it has proven possible to increase the selectivity, as shown also in the course of this thesis work.

# 2. PEPTIDOMIMETICS OF THE VIRAL TAT PROTEIN

## 2.1. Introduction

### 2.1.1. The importance of the Tat protein-TAR RNA interaction for viral transcription

The interaction between the viral transactivator protein Tat and the trans-activation responsive element (TAR) RNA plays a critical role in the replication of immunodeficiency viruses and has emerged in the last decade as an important target in antiviral therapy. The arginine-rich binding domain of the Tat protein binds to its cognate TAR RNA stem-loop structure located at the 5'-end of the viral mRNA, resulting in enhanced processivity of RNA polymerase II.

The non-phosphorylated RNA polymerase II (RNAPII) is recruited to the core promoter in proviral DNA, at the TATA box in the long terminal repeat (LTR), through its interaction with transcription factor IID (TFIID) and other components of the basal transcription apparatus, and so forms a pre-initiation complex (PIC). The carboxy-terminal domain (CTD) of the RNA polymerase II is then phosphorylated by cyclin-dependent kinase 7 (CDK7) found in TFIIH, and the modified polymerase clears then the promoter and begins the transcription of RNA (Figure 4).

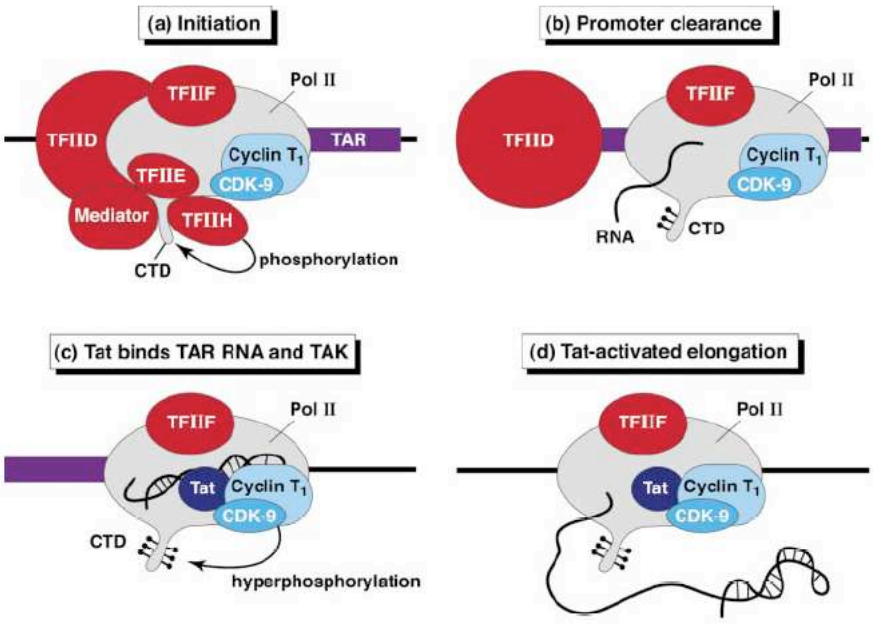



**Figure 4.** Model for activation of RNA polymerase II by Tat and cellular co-factors.[55]

Following promoter clearance, and before entering productive elongation, the processivity of RNAPII is controlled by the action of both negative (NELFs, DSIF) and positive elongation factors (P-TEFs). NELFs and DSIF have been proposed to be responsible for premature stopping and termination of the initiated polymerase, resulting in generation of short abortive transcripts. P-TEFs, such as p-TEFb, are known to mediate the transition from abortive into productive elongation by phosphorylating the CTD of polymerase II. P-TEFb is composed of two subunits, the catalytic subunit, cyclin dependent kinase CDK9, and the regulatory subunit, cyclin T1.

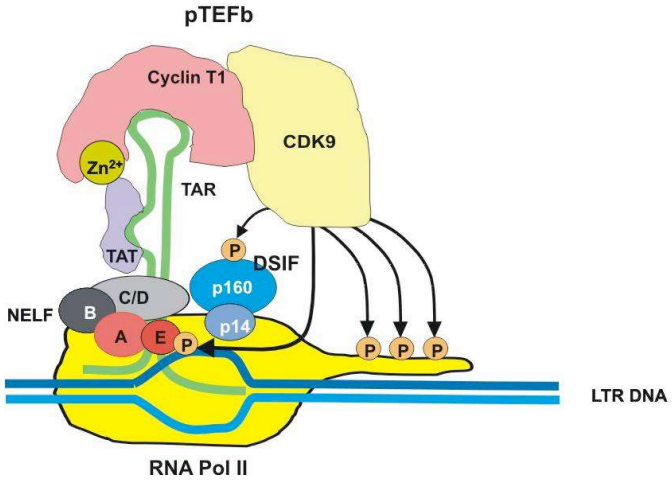



**Figure 5.** Transactivation mechanism of Tat protein. Only if the viral protein Tat is present and binds to the bulge of TAR the cellular transcription elongation factor p-TEFb is recruited. Upon phosphorylation of RNA polymerase II, NELF and DSIF by the kinase component CDK9 of p-TEFb, anti-termination takes place and TAR can be elongated to the full length viral RNA.

The complex of Tat with autophosporylated p-TEFb binds to the upper bulge and loop structures of the nascent RNA chain corresponding to the TAR RNA transcript, thereby positioning CDK9 to phosphorylate, **i**) the negative elongation factors, which repress transcription by binding to the lower stem in TAR, and **ii**) the CDT of RNAPII (hyperphosphorylation). These phosphorylation events result in the release of the transcriptional elongation block and the transcription elongation is promoted (Figure 5).[56,57]

The HIV-1 transactivator protein Tat consists of 101 amino acids in most HIV-1 isolates. There are four domains defined in the Tat protein sequence, an N-terminal proline rich/acidic region (amino acids 1-21), a region containing seven cysteines (22-37), a core region (38-48), a basic region enriched with arginine and lysine residues (49-57) that is highly conserved among different strains, and a cell-adhesion C-terminal region. Based on mutational analysis, Tat can be divided into two functional domains,

**i)** a functionally autonomous transcription-activation domain (amino acids 1-48), which is thought to be involved in metal ion binding, and **ii)** a conserved basic domain (47-59), essential for TAR recognition and binding. The basic domain is also important for translocation through the cell membrane.[58,59]

The Tat-mediated transactivation mechanism is a particularly attractive target for the development of new antiretroviral drug therapies, because Tat is required for viral gene expression not only during the exponential growth of the virus but also during the activation of the integrated proviral genomes that give rise to drug-resistant strains of HIV-1.[55]

### 2.1.2. A comparison between the HIV and BIV Tat transcriptional elements

Among the most studied RNA targets for the treatment of infectious diseases are the Tat RNA regulatory elements from HIV-1 and related viruses. The associated regulatory/accessory proteins are important features that differentiate lentiviruses from other retroviruses.

Since structural information about the HIV-1 Tat-TAR complex is not yet available, it was decided here to use the corresponding interaction from bovine immunodeficiency virus (BIV) to study TAR recognition by peptidomimetic ligands. Both viruses use similar transcriptional activation mechanisms and their Tat and TAR elements share structural and sequential similarities. Most importantly, the three-dimensional structure of BIV TAR in a complex with a Tat-derived peptide has been determined, revealing that BIV Tat forms a β-hairpin conformation in a complex with its cognate TAR RNA (Figure 6). These observations make BIV an ideal model for studying this interaction, with the possibility to export the information derived from BIV into the HIV system.

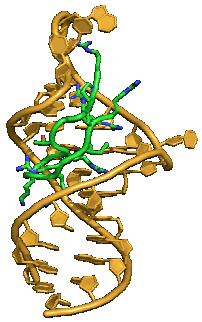


**Figure 6.** The BIV Tat peptide (green) adopts a β-hairpin conformation upon binding to TAR RNA (PDB 1MNB).

BIV, first isolated in 1969, causes lymphocytosis, lymphadenopathy, progressive weakness, and central nervous system disorders in infected cattle. It belongs to the unique genus *Lentivirinae*, which share structural, genetic, biological and/or pathological properties and infect monocyte/macrophage cells. BIV is the most complex of non-primate lentiviruses and bears a strong resemblance to primate viruses including HIV-1 and HIV-2, as well as simian immunodeficiency virus (SIV). Besides genes encoding the obligate retroviral Gag, Pol, and Env polyproteins, BIV encodes six non-structural/accessory proteins, which are called Tat, Rev, Vif, Vpy, Vpw, and Tmx. These proteins play a critical role in the viral replicative cycle and contribute significantly to the pathogenesis of BIV.[60]

BIV Tat protein contains five structural domains similar to those in primate lentiviruses, an acidic amino terminal, a cysteine-rich, a highly conserved core, a basic and a carboxy-terminal region. The functional domains include an activation domain and an RNA-binding domain.

BIV Tat protein residues directly interacting with the RNA form a basic, arginine-rich domain, and include Arg70, Gly71, Thr72, Arg73, Arg77 and Ile79, while the Gly74 and Gly76 play an important conformational role (Figure 7). The BIV Tat-derived peptide adopts a β-hairpin structure defined by an antiparallel β-sheet. The turn of the β-hairpin occurs between Arg73 and Arg77, while two Gly residues (74 and 76) facilitate the turn. The Tat peptide penetrates deep into the major groove and the tight fit results in substantial widening of the groove (from approximately 10Å to 16Å).[26,61,62]

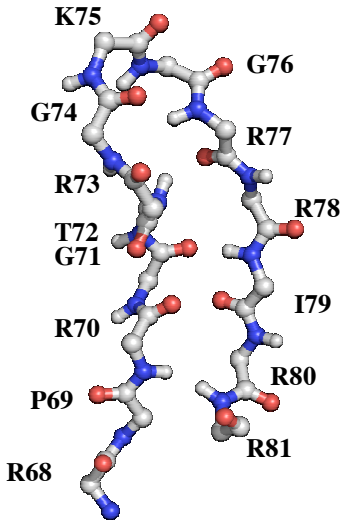



**Figure 7.** RNA-binding domain of BIV Tat protein in a β-hairpin conformation.

Binding to HIV-1 TAR RNA is mediated by residues 49-57 in the HIV-1 Tat protein. Mutagenesis studies showed that a single arginine at position 52 makes the key contribution to the binding specificity (Figure 8).[63] Studies with isolated peptides have shown that additional basic amino acids that surround this arginine help enhance the affinity and kinetic stability of the complex as well as improving binding specificity.[64]

The arginine-rich domain of HIV Tat, like that of BIV, does not appear to adopt a stable structure in solution. Mutational studies suggest that the bound form of Tat is might be helical, but the actual bound conformation has so far not been proven.[65]

BIV Tat $^{65}$SGPRP**RGTR**GKG**R**R**I**RR$^{81}$

HIV Tat $^{44}$GISYGRKK**R**QR**R**RPPQGS$^{62}$

**Figure 8.** Comparison of HIV-1 Tat and BIV Tat arginine-rich RNA-binding domains with the amino acids critical for binding TAR RNA in red.

The TAR element varies in length and structure among lentiviruses. The BIV and HIV-1 TAR RNA form a stem-bulge-loop hairpin structure composed of 28 and 59 nucleotides, respectively, which comprises the Tat binding site (Figure 9). Tat interacts in BIV transactivation with nucleotides G11-C25, G14-C23 and C15-G22 located in the stem of TAR, and it directly binds the bulge at U10. Then a base-triple RNA structure comprised of U10-A13-U24 is formed and is consolidated by hydrogen bonds.[66,67] The central loop of BIV TAR RNA, comprising nucleotides CAUU, is not directly involved in the binding of BIV Tat, contrary to the situation in HIV.[68,69]

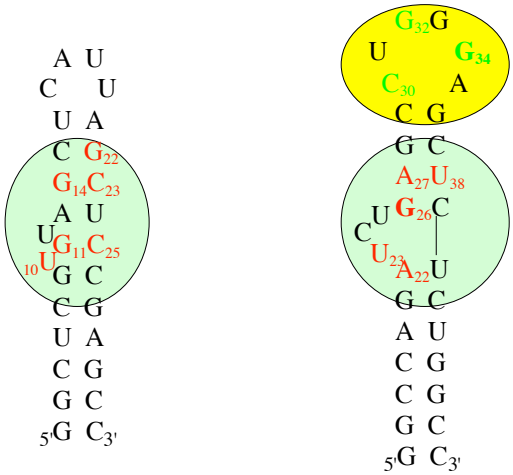



**Figure 9.** Comparison of the BIV and HIV-1 TAR RNA sequences. The nucleotides in red play a critical role in Tat recognition, the nucleotides in green for HIV-1 are critical for cyclin T1 binding.

Despite extensive efforts, a high-resolution structure of the HIV TAR-Tat complex has not yet been obtained. However, the NMR solution structures of the HIV TAR RNA and the Tat protein have been solved individually. Several NMR studies have investigated the HIV TAR RNA structural change that takes place upon binding to arginine[70] (Figure 10), argininamide[71] or Tat-derived peptides[72,73] and they gave deeper insights into the principles of TAR RNA recognition by Tat. Together with extensive mutagenesis, chemical probing and peptide binding studies, the key elements required for TAR recognition by Tat have been defined, and it has been shown that the Tat binding site surrounds a UCU bulge located near the apex of TAR. In detail, it has been shown that a critical role in TAR recognition is played by U23, which forms a base triple with A27 and U38, and stacks on A22 and along G26. In addition, several phosphate groups adjacent to the bulge, including P21 (between G21 and A22), P22, and P40, are suggested to be essential for RNA recognition by HIV Tat.

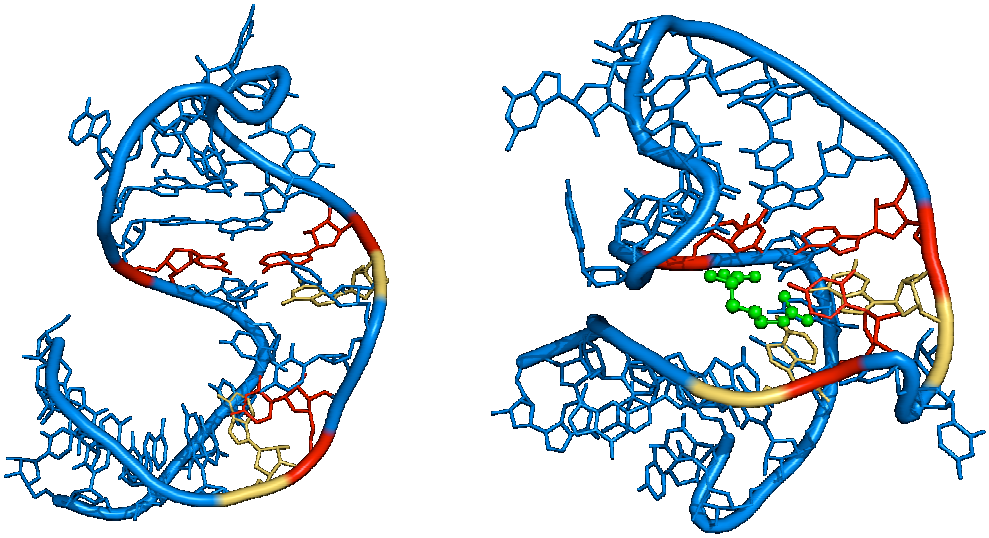


**Figure 10.** Major groove view of free HIV-1 TAR RNA (left, PBD 1ANR), and TAR-arginine complex (right, PDB 1ARJ). The nucleotides critical for binding are highlighted, in red U23, A27 and U38, which form a base triple, and in yellow A22 and G26. Arginine is shown in green.

In both cases, HIV-1 and BIV, TAR-Tat complex formation induces global RNA conformational changes and alters the internal dynamics of the nucleotides surrounding the binding region in the TAR RNA. The bulge nucleotides help in widening of the major groove of the RNA, which is normally very narrow (~11Å) and resembles the minor groove of DNA in its interactions with proteins.

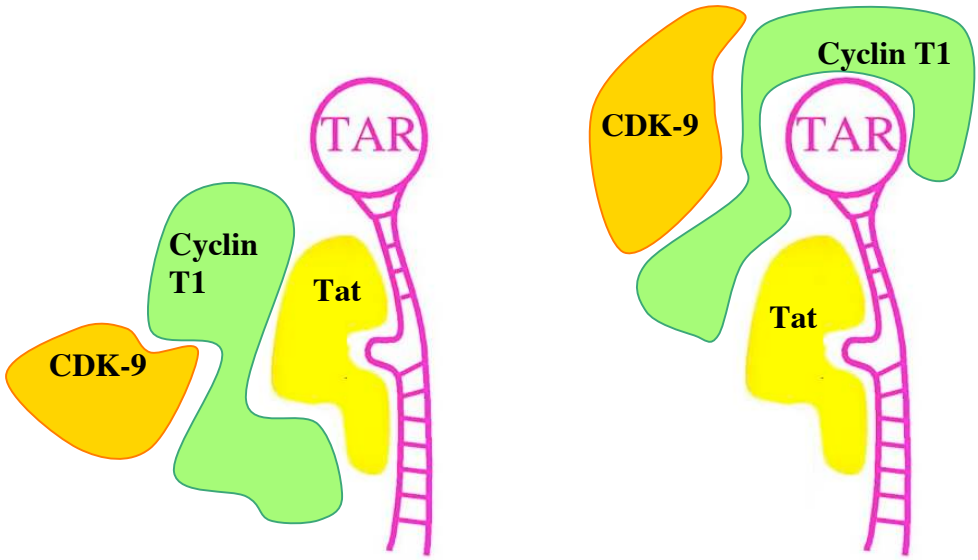
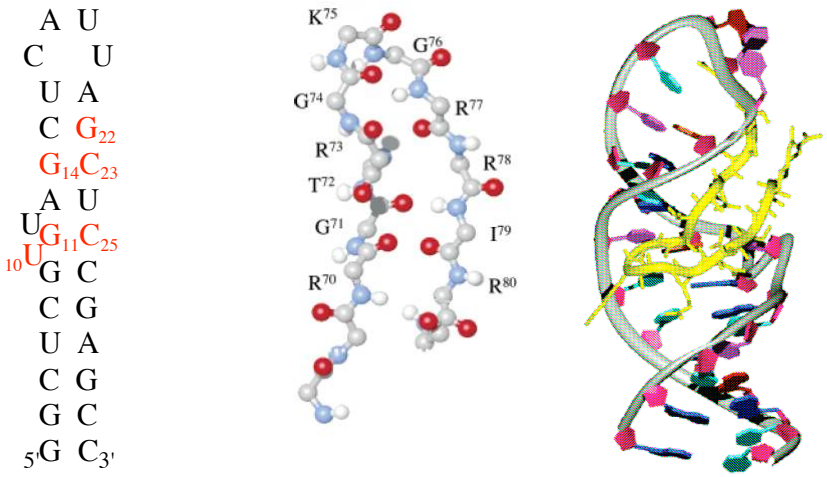



**Figure 11.** Schematic model of cyclin T1-Tat-TAR interaction shown on the left for BIV and right for HIV. Binding of HIV Tat and cyclin T1 to HIV TAR RNA is highly cooperative, while for BIV it occurs independently.[74]

Although the TAR-dependent transactivation mechanism shares similarities among lentiviruses, there are differences in the Tat-pTEFb-TAR recognition event and the transactivation complex formation. HIV-1 Tat binds directly through its transactivation domain (amino acids 1-48) to the cyclin subunit of the pTEFb (cyclin T1), and induces loop sequence specific binding of this cyclin to TAR RNA.[75] Cyclin T1 and Tat binding to HIV TAR RNA is highly cooperative. Cyclin T1 does not bind to TAR in the absence of Tat, and the HIV Tat RNA-binding domain binds TAR at least an order of magnitude more weakly in the absence of cyclin T1.[72] The cyclin T1-Tat heterodimer directly binds to TAR RNA in the U-rich RNA budge, and this binding facilities the interactions of the cyclin T1-TAR RNA in the loop region (Figure 11, right).[76] The interaction of HIV Tat with the human cyclin T1, increases greatly the affinity and specificity of the binding between HIV Tat and TAR. In sharp contrast, the BIV Tat may recognize the BIV TAR with high affinity in the presence or absence of cyclin T1 (Figure 11, left). Consequently, the BIV Tat-TAR interaction takes place in most mammalian cells, including murine, canine, rabbit and human cells, indicating flexibility for BIV Tat to recruit cyclin T1.[68,74]

### 2.1.3. Structure of the BIV Tat-TAR RNA complex solved by NMR

Two decades of extensive studies on the Tat-TAR interaction have given deep insights into the principles of protein-RNA recognition and the transactivation mechanism in immunodeficiency lentiviruses. So far, only the structure of a BIV Tat derived peptide-TAR complex has been solved. In 1995 two independent groups determined NMR-structures of the BIV Tat RNA binding domain (residues 68-81 in one study[26] and 65-81 in the other[62]) in complex with the BIV TAR RNA stem-bulge-loop structure. Although using a different number of distance restraints, the two structures are globally in agreement, and the specific interactions seen in the complex explain the available *in vitro* and *in vivo* mutagenesis data. Both groups used a 28-nucleotide fragment of BIV TAR RNA and the arginine-rich RNA binding domain of BIV Tat RNA (Figure 12).




**BIV Tat**: $^{65}$SGPRP**RGTR**GKG**R**R**I**RR$^{81}$

**Figure 12.** Secondary structures of the BIV Tat and TAR RNA binding regions. Left: BIV TAR stem-loop structure; Middle*:* arginine-rich TAR-binding domain of BIV Tat protein in a β-hairpin conformation. Nucleotides and amino acids important for binding are highlighted in both structures. Right: the three-dimensional structure of the complex between BIV Tat and TAR RNA (PDB 1MNB).

Upon binding, the BIV Tat peptide, which is unstructured in aqueous solution, adopts a β-hairpin conformation and penetrates the widened (from ~10Å to 16Å) major groove of the RNA specifically (Figure 12, right). The β-hairpin structure is defined by an antiparallel β-sheet between the G71-T72-R73 and R77-R78-I79 segments. The β-turn of type I is defined by Gly74 to Arg77 with a Gly76 promoting turn formation.

The formation of the base triple (U10-A13-U24) in the BIV Tat-TAR complex facilitates both the deformation of the backbone and the orientation of the phosphate between G9 and U10. This phosphate participates in a key intermolecular interaction with an Arg73 side chain. The guanidinium group forms hydrogen bonds with both the major groove edge of G11 and the phosphate oxygen atom. The aromatic ring of U10, positioned by formation of the base triple, interacts via van der Waals contacts with Ile79, whose side chain is packed into the major groove edge of the base triple. Thr72 is also involved in critical contacts with the ribose of G22. A hydrogen bond to the RNA backbone phosphate between G22 and C23 was suggested by Puglisi,[26] whereas hydrogen bonds between the carbonyl oxygen and the amino group of C23, and hydroxyl oxygen and phosphate between C23 and U24 were suggested by Patel.[62] The guanidinium groups of Arg70 and Arg77 make contacts with guanines at the major-groove edge, Arg70-G14 via two hydrogen bonds and Arg77-G9 through two bifurcated hydrogen bonds. The intermolecular hydrogen bond between the amide proton of Gly71 and G22 leaves no room for any larger side chains, as it would point directly into A21. The G74 and G76 have no significance for binding to RNA but cannot be mutated without loss of activity, due to the steric hindrance that larger side chains would cause.

### 2.1.4. Inhibition of TAR RNA-Tat protein interaction

Among the earliest findings of ligands that bind to TAR RNA and prevent Tat TAR complex formation were the well known natural products, the aminoglycosides, already discussed in detail in paragraph 1.1.3. However, most aminoglycosides bind to a variety of RNA targets with lack of selectivity, which often results in severe toxicity, thus making them unattractive for drug development.

During the last decade, a limited number of non-peptidic small-molecule ligands (following Lipinski's rule, MW<500) have been reported to inhibit Tat-TAR complex formation. Many groups focused on the development of antagonists that posses a high and specific affinity for the bulge region of TAR RNA.

By high-throughput *in vitro* screening of 150'000 compounds, Mei et al.[77] identified two promising candidates, **2,3-dioxo-8- [2-(5-tetrazolyl)]-2,3,4,7,8,9-hexahydro-1*H*-6-nitrocyclopenta[*f*]quinoxaline (5)** and **2,4,5,6-**

**tetraaminoquinozaline (6).** Studying their interaction with TAR RNA it was discovered that they bind to two different binding sites. Quinoxaline interacts with the bulge region, i.e. the Tat binding site, while tetraaminoquinozaline binds to the loop of the RNA, the cyclin T1 binding site.[78] A promising class of 6-aminoquinolone TAR ligands, reported by Cecchetti et al.[79] was discovered by testing antimicrobial quinolone DNA gyrase inhibitors. The lead compound, a **4-(2-pyridyl)-1-piperazinyl-substituted 6-amino-quinolone (WM5) (7)** exhibited submicromolar affinity to TAR RNA. WM5 was able to inhibit Tat-TAR complex formation with an apparent $K_i$ of 3.5 μM and exhibited inhibitory potency of HIV replication in infected cells in low micromolar range. It has been proven that this compound binds to the bulge region of the wild-type TAR RNA.[80]

Well established *in silico* techniques were employed by the James' group[81] to predict potent Tat-TAR inhibitors, which then were tested *in vitro*. Apart from the ligands that were already known, e.g. aminoglycosides, many novel structures, such as **anthraquinone (8)** or **trisamine-derivatives (9)** were identified and exhibited an $IC_{50}$ value of approximately 1μM. The same group, in 2002,[82] using computational screening of a 181,000-compound available chemicals directory against the three-dimensional structure of TAR predicted 500 potential TAR ligands, from which 50 were tested *in vitro*. Among these compounds, the most promising lead was **acetylpromazine (10)** (also known as acetopromazine or acepromazine), a phenothiazine derivative, which represents a new class of TAR ligands with ~270μM binding affinity, good bioavailability and low toxicity. NMR studies[83] revealed that this compound binds to the bulge region of TAR, where the planar heteroaromatic ring system stacks between base pairs G26-C39 and A22-U40. It is noteworthy that acetylpromazine structurally resembles synthetic **aminoacridine (11)**, a ligand of TAR that has been developed by Novartis Pharma using a rational design approach. These so called "In-PriNts", inhibitors of protein-ribonucleotide sequences, should be tripartite and consist of **i)** a polyaromatic/heterocyclic moiety, with the potential for a stacking interaction inside the cavity of TAR, delimited by U23 as well as by bases below and above the bulge (A22 and G26, respectively), **ii)** a feature providing positive charges for interaction with the phosphate backbone of RNA, and **iii)** a spacer to connect the stacking entity with the RNA binding part.[84] A similar rational approach was taken by Yuan et al.[85] and resulted in the design of **substituted purines (12)** containing a side

chain with a terminal amino or guanidyl group. All the compounds effectively block TAR transactivation in human 293T cells, show low cytotoxity and interact with the bulge region of TAR.

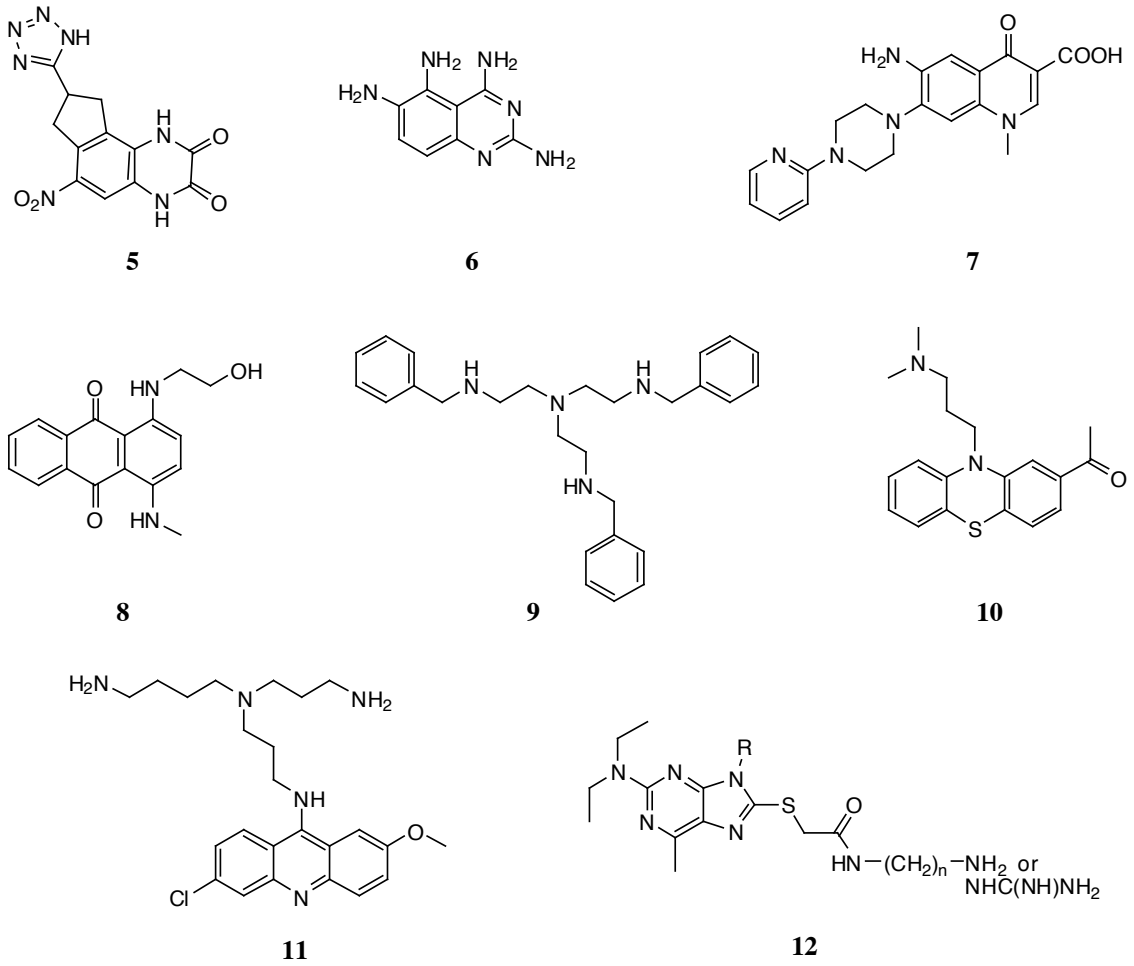



**Figure 13.** Tat-TAR interaction inhibitors: (5) derivative of quinoxaline-2,3-dione, (6) tetraaminoquinozaline, (7) WM5, (8) anthraquinone, (9) trisamine-derivatives, (10) acetylpromazine, (11) substituted aminoacridine, (12) substituted purine.

Peptides and peptidomimetics that specifically target RNA structures are potential antagonists of RNA-binding proteins and are therefore capable of controlling cellular functions. During the last decade, several peptidic structures have been developed that specifically interact with the bulge region of TAR. The starting point for many studies was the Tat sequence that is responsible for binding to TAR RNA. Tat-derived **oligocarbamate (13)** and **oligourea (14)** contain basic side chains like those in the arginine-rich region of full length Tat protein and show inhibition in the low μM range.[86]

A combinatorial library of tripeptides of L- and D- natural amino acids was prepared by Hwang et al.[87] Out of 24389 variations, eight sequences were found promising, among them two exhibit submicromolar affinity towards TAR RNA and inhibit transcriptional activation by Tat protein in human cells with an $IC_{50}$ of ≈50 nM. The best binding tripeptide **(15)** was then incorporated into the cyclic structure **(16)** to improve pharmacokinetic properties, which showed increased inhibitory activity.[88]

Another peptoid/peptide **(17)** was presented by Hamy et al.[89] that was thought to inhibit the Tat/TAR association and block HIV-1 replication in primary human lymphocytes. However, later studies showed this whole cell inhibitory activity arises by a mechanism unrelated to Tat-TAR inhibition.[90,91]

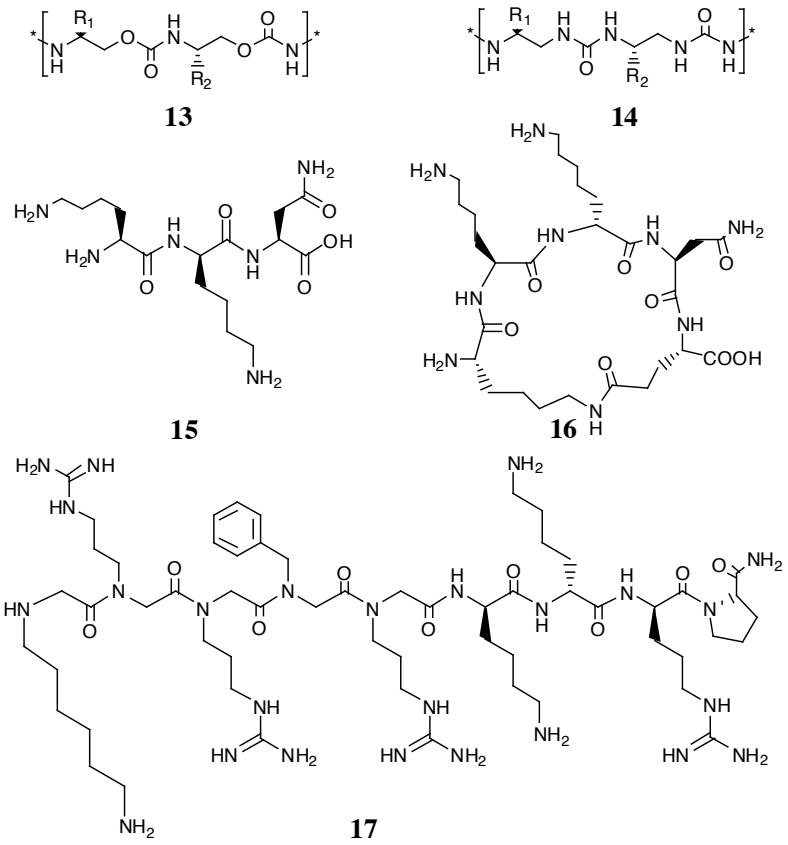



**Figure 14.** Peptidic inhibitors of the Tat-TAR interaction: (13) Tat-derived oligocarbamate, (14) Tat-derived oligourea, (15) tripeptide H-Lys-D-Lys-Asn-OH, (16) pentapeptide cyclo(Lys-D-Lys-Asn-Asp-ε-Lys), (17) CGP64222.

### 2.1.5. Previous work

On the basis of the structural information obtained for the BIV Tat/TAR RNA presented a decade ago by Puglisi et al.[26] followed by many further studies, it was possible to design the first peptidomimetic Tat-TAR interaction inhibitors. The first attempt to use the critical binding sequence of wild type Tat peptide was reported by Runyon.[92] A cyclic 14-amino acid peptide derived from the arginine-rich BIV Tat motif, cyclo(dYGRGTRGKGRRIVN), showed a similar mode of action and almost the same affinity as the linear wild-type peptide.

Following this observation, Athanassiou et al.[54] reported a library of 8 cyclic peptidomimetics attached to a β-hairpin-inducing template. The starting point was a mimetic obtained by transplanting the residues 70-80 from BIV Tat onto a D-Pro-L-Pro template. In order to form a regular 2:2 β-hairpin conformation, the loop was extended to 12 amino acids, and the critical amino acids for TAR binding were kept constant (Figure 15).

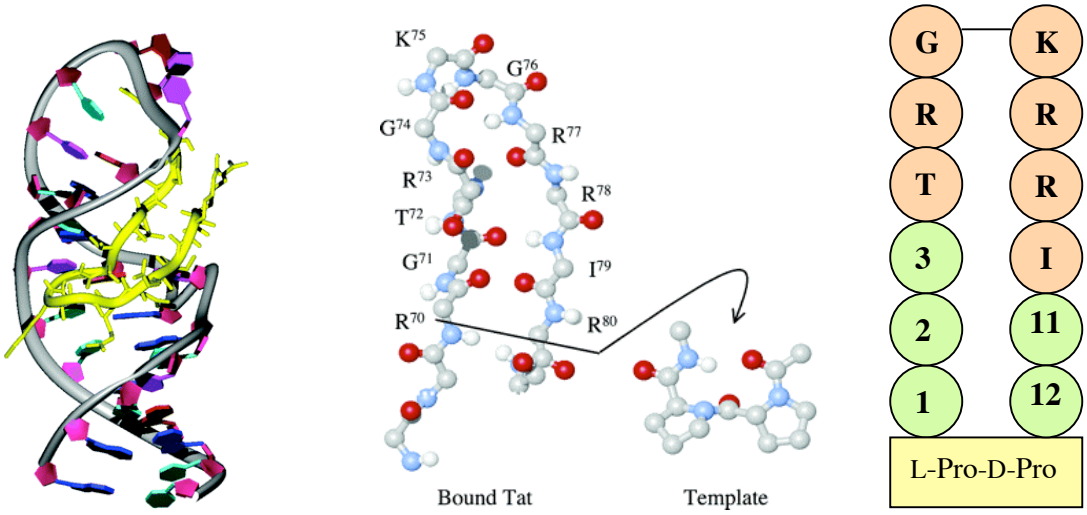



**Figure 15.** The design of BIV Tat mimetics. Left: the NMR structure of BIV Tat-TAR complex; middle: the conformation of the Tat bound to TAR RNA, moulded onto a D-Pro-L-Pro template; right: the representation of a typical 2:2 β-hairpin peptidomimetic. The residues 4-10 remain constant.

A set of cyclic peptidomimetics grafted onto a D-Pro-L-Pro template gave very promising results, summarized in Table 3.

| Mimetic | Position | | | | | | | | | | | | $K_d$ [μM] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* | |
| **BIV0** | - | R | G | T | R | G | K | G | R | R | I | R | 40 |
| **BIV1** | I | R | G | T | R | G | K | R | R | I | R | V | 30 |
| **BIV2** | R | V | R | T | R | G | K | R | R | I | R | V | 0.150 |
| **BIV3** | I | Y | R | T | R | G | K | R | R | I | R | T | nb |
| **BIV4** | Y | R | G | T | R | G | K | R | R | I | Y | V | >50 |
| **BIV5** | R | R | G | T | R | G | K | R | R | I | G | R | 1-2 |
| **BIV6** | V | R | G | T | R | G | K | R | R | I | K | Y | >50 |
| **BIV7** | V | R | R | T | R | G | K | R | R | I | K | Y | nb |
| **BIV8** | K | R | G | T | R | G | K | R | R | I | G | Y | >50 |

**Table 3.** Sequences of β-hairpin peptidomimetics. Positions 1-12 refer to residues 1-12 as mounted on the D-Pro-L-Pro template (Figure 15, right). The dissociation constants ($K_d$) were determined for the interaction of each mimetic with BIV TAR$^{4-31}$ RNA (EMSA) (nb=no binding detected). The 17-residue wild-type peptide Ac-$^{65}$SGPRPRGTRGKGRRIRR$^{81}$-$NH_2$ (BIV Tat (65-81)) was used as a positive control in all binding studies, with a $K_d$ value of 50 nM.

Among the peptidomimetics, BIV2 showed a very high affinity to TAR RNA and proved to bind to it with high specificity, since the EMSA was performed in the presence of a large excess of tRNA [for details see Paragraph 2.2.3].

In order to identify the amino acids important for the binding, each position of BIV2, except Gly at position 6, was mutated to Ala. Eleven new peptidomimetics, BIV2A1-BIV2A12, were produced in this way and assayed. The mutation of many amino acids to Ala was found to have a large detrimental effect on the binding (Table 4).

**Table 4.** Alanine scan of the BIV2 peptidomimetic. $K_d$ values determined by EMSA (nb=no binding detected).

| Mimetic | Position | | | | | | | | | | | | $K_d$ [μM] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* | |
| **BIV2** | R | V | R | T | R | G | K | R | R | I | R | V | 0.150 |
| **BIV2A1** | **A** | V | R | T | R | G | K | R | R | I | R | V | nb |
| **BIV2A2** | R | **A** | R | T | R | G | K | R | R | I | R | V | 0.3 |
| **BIV2A3** | R | V | **A** | T | R | G | K | R | R | I | R | V | nb |
| **BIV2A4** | R | V | R | **A** | R | G | K | R | R | I | R | V | 0.3 |
| **BIV2A5** | R | V | R | T | **A** | G | K | R | R | I | R | V | nb |
| **BIV2A7** | R | V | R | T | R | G | **A** | R | R | I | R | V | nb |
| **BIV2A8** | R | V | R | T | R | G | K | **A** | R | I | R | V | nb |
| **BIV2A9** | R | V | R | T | R | G | K | R | **A** | I | R | V | nb |
| **BIV2A10** | R | V | R | T | R | G | K | R | R | **A** | R | V | 5 |
| **BIV2A11** | R | V | R | T | R | G | K | R | R | I | **A** | V | nb |
| **BIV2A12** | R | V | R | T | R | G | K | R | R | I | R | **A** | 5 |

Furthermore, NMR spectroscopic studies on BIV2 free in solution revealed a stable β-hairpin conformation for this peptide and this feature facilitated the structure elucidation of the BIV2-TAR RNA complex. BIV2 induces a conformational change in TAR leading to a structure very similar to that seen in Tat-bound TAR RNA. The structure of the complex between the BIV2 peptide and BIV TAR RNA determined by NMR spectroscopy revealed that BIV2 binds to RNA in an orientation that is flipped upside-down compared to that expected based on a comparison to the wild-type Tat peptide, but the key intermolecular contacts that direct the RNA conformational change are maintained (Figure 16).[93]

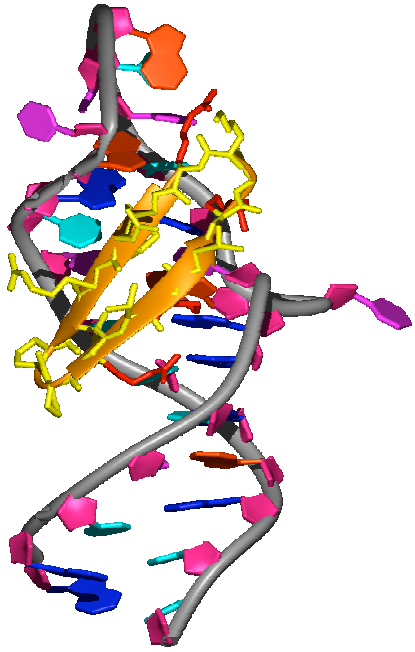


**Figure 16.** The structure of the complex between the BIV2 cyclic peptide and BIV TAR RNA determined by NMR spectroscopy.[93]

### 2.1.6. Project outline

The successful design of the small library of cyclic Tat peptidomimetics that included a relatively potent inhibitor of BIV Tat-TAR, called BIV2, and determination of the structure of the complex of this ligand bound to BIV TAR opened the way for optimisation and further structural studies. On the basis of these findings, the purpose of this work was to optimise the lead compound, BIV2, and by doing so, to gain new insights into the origins of the high affinity and specificity of this ligand-RNA interaction. It was assumed that by introducing single mutations into BIV2 it would be possible to identify alternative side chains that would provide optimal contacts to the RNA, and that by combining individual changes and incorporating multiple mutations it would be possible to increase affinity by additive effects. Additionally, to evaluate if the interactions observed are specific, the peptidomimetics should be assayed with other related RNAs, namely HIV-1 TAR RNA and RRE RNA.

Furthermore, since the transactivation mechanism for BIV and HIV shares many similarities, it seemed likely that the HIV Tat-TAR system could be mimicked by the BIV complex and the information derived from studies on BIV Tat inhibitors could be exported, at least in part, to derive inhibitors of the human virus. Therefore, all the previously prepared 14-amino acid BIV Tat peptidomimetics were assayed with HIV-1 TAR RNA in order to identify lead compounds. Additionally, it was envisioned that introducing an additional four amino acids and increasing the number of residues in the sequence from 14 to 18 might enable the ligand to interact not only with the major groove, but additionally with the tip of the loop of the TAR RNA. This might then lead to inhibition of the cyclin T1-TAR RNA interaction, which is known to be necessary for effective transactivation. Moreover, it was reasoned that introducing a disulfide bridge into these 18-amino acid peptides might stabilize the β-hairpin conformation.

The ultimate target was to evaluate the ability of the selected peptides to cross the cell membrane, to test their stability in biological settings, and to prove their activity using *in vivo* assays.

## 2.2. Results

### 2.2.1. BIV Tat peptidomimetics

- **LIBRARY L1: Attempts to identify side chains optimal for RNA contacts by single mutations of the residues in BIV2.**

The sequence of BIV2, a sub-micromolar ligand ($K_d$=150 nM) of BIV TAR RNA with a stable β-hairpin conformation both free in solution and when bound to RNA, served as a lead for further optimisation. The β-hairpin structure of BIV2 bound to RNA allows us to identify the side chains interacting with the major groove (Arg1, Arg3, Arg5, Arg8, Ile10, Val12) and the residues exposed to the solvent (Val2, Thr4, Arg9, Arg11). The other two amino acids, Gly6 and Lys7, are at the tip of the β-hairpin and point toward the tip of the RNA.[93]

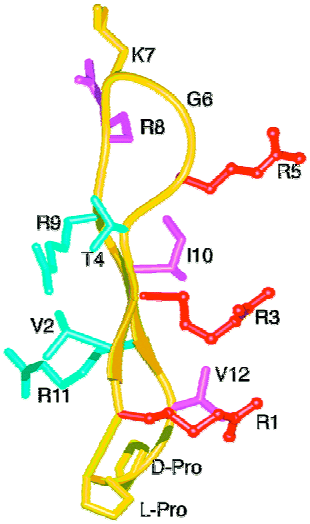



**Figure 17.** β-Hairpin structure of BIV2 [cyclo-(D-Pro-L-Pro-Arg-Val-Arg-Thr-Arg-Gly-Lys-Arg-Arg-Ile-Arg-Val)] bound to BIV TAR RNA. pink, red: the residues interacting with the major groove of RNA, green: the residues exposed to the solvent. The tip of the β-hairpin is at the top and the template D-Pro-L-Pro at the bottom.

In the first round of optimisation (Table 5), positions in the RNA-contacting face of BIV2 were targeted for mutation in an effort to identify alternative side chains that would provide optimal contacts to the RNA (see Figure 17). Arginine residues at positions 1 and 3, found from the complex structure to have hydrogen bonding interactions with TAR, were initially kept constant, with the exception of Arg3Lys, Arg3Cit, Arg3Orn to deliver more information on the nature of the Arg3 interaction

with the RNA. Arg5 and Arg8 were more freely mutated, as in the complex they show cation-π interactions rather than hydrogen bonding as observed in Arg1 and Arg3. Mutations of these residues to other charged amino acids would help to distinguish between the requirements for hydrogen bonding and electrostatic interactions. Hydrophobic residues (positions 10 and 12) were mutated to other hydrophobic or charged residues in an effort to optimise the intermolecular contacts at these positions. Positions 2, 4, 9, 11 on the solvent exposed face were also mutated in an effort to further increase the affinity by indirect conformational effects and to reduce the overall charge of the peptide. The residues at the tip of the β-hairpin were kept constant, as they were thought to be important for defining the β-turn and formation of a stable hairpin conformation.

**Table 5.** Library L1 of BIV Tat mimetics. Residues 1 and 12 are attached to the D-Pro-L-Pro template. $K_d$ values [μM] were determined by EMSA (nb=no binding detected).

| **Mimetic** | **Position** | | | | | | | | | | | | **$K_d$ BIV** | **$K_d$ HIV** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | | |
| **BIV2** | **R** | **V** | **R** | **T** | **R** | **G** | **K** | **R** | **R** | **I** | **R** | **V** | 0.15 | >5 |
| **L1-01** | R | V | R | T | R | G | K | R | R | I | R | **L** | 0.15 | 0.5 |
| **L1-02** | R | V | R | T | R | G | K | R | R | I | R | **I** | 0.05 | 0.5 |
| **L1-03** | R | V | R | T | R | G | K | R | R | I | R | **F** | 0.15 | 0.5 |
| **L1-04** | R | V | R | T | R | G | K | R | R | I | R | **T** | 0.5 | 0.5 |
| **L1-05** | R | V | R | T | R | G | K | R | R | I | R | **N** | 0.5 | 0.15 |
| **L1-06** | R | V | R | T | R | G | K | R | R | I | R | **R** | 0.15 | 0.15 |
| **L1-07** | R | **T** | R | T | R | G | K | R | R | I | R | V | 0.05 | 0.15 |
| **L1-08** | R | **N** | R | T | R | G | K | R | R | I | R | V | 0.15 | 0.25 |
| **L1-09** | R | V | R | T | R | G | K | R | R | **L** | R | V | 0.5 | >10 |
| **L1-10** | R | V | R | T | R | G | K | R | R | **Cha** | R | V | 5 | >10 |
| **L1-11** | R | V | R | T | R | G | K | R | R | **F** | R | V | 0.5 | >5 |
| **L1-12** | R | V | R | T | R | G | K | R | R | **Y** | R | V | 0.5 | 1 |
| **L1-13** | R | V | R | T | R | G | K | R | R | **N** | R | V | 0.25 | 0.5 |
| **L1-14** | R | V | R | T | R | G | K | R | R | **Q** | R | V | 0.5 | 0.5 |
| **L1-15** | R | V | R | **Q** | R | G | K | R | R | I | R | V | 0.05 | 0.5 |
| **L1-16** | R | V | R | **V** | R | G | K | R | R | I | R | V | 0.1 | 0.5 |
| **L1-17** | R | V | R | **Y** | R | G | K | R | R | I | R | V | 0.1 | 1 |
| **L1-18** | R | V | R | **K** | R | G | K | R | R | I | R | V | 0.5 | 0.5 |
| **L1-19** | R | V | **K** | T | R | G | K | R | R | I | R | V | >5 | 5 |
| **L1-20** | R | V | **U** | T | R | G | K | R | R | I | R | V | >5 | >10 |
| **L1-21** | R | V | **O** | T | R | G | K | R | R | I | R | V | 1.5 | 1 |
| **L1-22** | R | V | R | T | **K** | G | K | R | R | I | R | V | >5 | >10 |
| **L1-23** | R | V | R | T | **U** | G | K | R | R | I | R | V | >5 | >10 |
| **L1-24** | R | V | R | T | **O** | G | K | R | R | I | R | V | >5 | 2.5 |
| **L1-25** | R | V | R | T | **N** | G | K | R | R | I | R | V | >5 | >10 |
| **L1-26** | R | V | R | T | **Q** | G | K | R | R | I | R | V | >5 | >10 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **L1-27** | R | V | R | T | **I** | G | K | R | R | I | R | V | >5 | >10 |
| **L1-28** | R | V | R | T | **L** | G | K | R | R | I | R | V | >5 | nb |
| **L1-29** | R | V | R | T | **Y** | G | K | R | R | I | R | V | >5 | nb |
| **L1-30** | R | V | R | T | R | G | K | **K** | R | I | R | V | 2.5 | 2.5 |
| **L1-31** | R | V | R | T | R | G | K | **U** | R | I | R | V | >5 | nb |
| **L1-32** | R | V | R | T | R | G | K | **O** | R | I | R | V | 2.5 | 0.5 |
| **L1-33** | R | V | R | T | R | G | K | **N** | R | I | R | V | >5 | nb |
| **L1-34** | R | V | R | T | R | G | K | **F** | R | I | R | V | >5 | nb |
| **L1-35** | R | V | R | T | R | G | K | **Y** | R | I | R | V | >5 | >10 |
| **L1-36** | R | V | R | T | R | G | K | R | **K** | I | R | V | 0.15 | 0.25 |
| **L1-37** | R | V | R | T | R | G | K | R | **U** | I | R | V | nb | nb |
| **L1-38** | R | V | R | T | R | G | K | R | **O** | I | R | V | 0.25 | 0.25 |
| **L1-39** | R | V | R | T | R | G | K | R | **N** | I | R | V | >5 | nb |
| **L1-40** | R | V | R | T | R | G | K | R | **Q** | I | R | V | >5 | nb |
| **L1-41** | R | V | R | T | R | G | K | R | R | I | **K** | V | 0.15 | 0.25 |
| **L1-42** | R | V | R | T | R | G | K | R | R | I | **U** | V | >5 | nb |
| **L1-43** | R | V | R | T | R | G | K | R | R | I | **O** | V | 0.1 | 0.25 |
| **L1-44** | R | V | R | T | R | G | K | R | R | I | **N** | V | 5 | nb |
| **L1-45** | R | V | R | T | R | G | K | R | R | I | **Q** | V | >5 | nb |

### 2.2.2. Synthesis of Tat peptidomimetics

All the Tat peptidomimetics described in this work were synthesised according to a mixed solid-phase solution-phase strategy developed earlier.[94] Briefly, the first amino acid was coupled to freshly activated, acid-labile 2-chlorotrityl chloride resin with 4-equivalents DIEA in DCM. Generally, Fmoc-L-Pro-OH was chosen as the first amino acid to be attached to the resin, since this allowed for the synthesis of larger batches of preloaded resin that could be used for the synthesis of almost all the peptidomimetics. The linear, fully protected precursors were synthesised employing standard Fmoc-chemistry on a MultisynTech Syro II peptide synthesizer (parallel synthesis up to 48 peptides) on ~0.07 mmol scale or, when larger quantities were required, on an Applied Biosystems 433A peptide synthesizer on a 0.25 mmol scale. Generally, 4 equivalents of the appropriately side-chain protected N-Fmoc amino acid are activated with HBTU, HOBt (4 eq each) and DIEA (8 eq) in DMF. In the parallel synthesis every coupling was performed twice. Fmoc deprotection was carried out in 20-40% piperidine in DMF solution.

After completion of the peptide chain assembly, the linear peptides were cleaved from the resin with 0.8% TFA in DCM. The crude linear peptides were cyclized in DMF using HATU, HOAt and DIEA in DMF. The final deprotection was achieved with TFA:$H_2O$:TIPS (95:2.5:2.5) cocktail and the crude cyclic peptides were

precipitated with ice-cold diethyl ether (Figure 18). The purification of the peptide libraries was performed by preparative RP-HPLC using $H_2O$+0.1% TFA and MeCN+0.1% TFA and yielded the desired peptides of ≥95% purity and in 20-35% yield. All the peptides were characterised by analytical HPLC and LC-MS (see Appendix 2 for the analytical data).

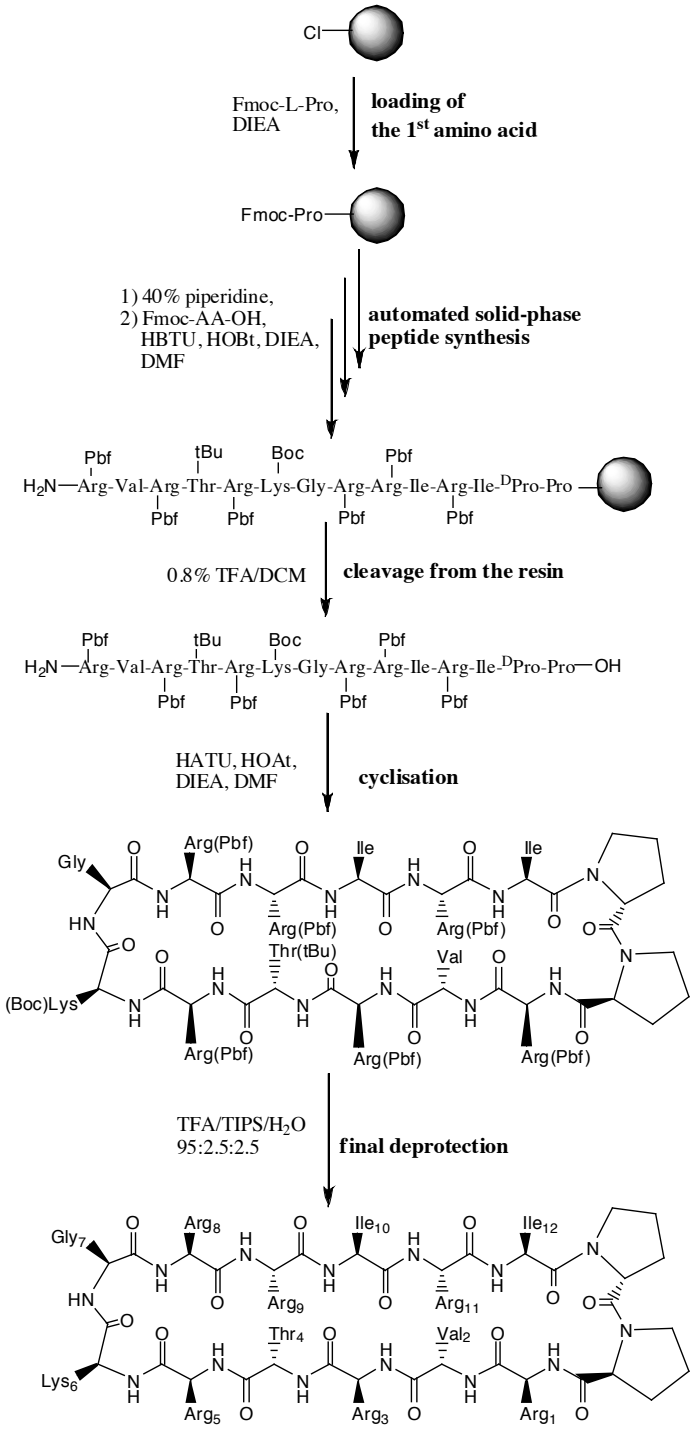



**Figure 18.** Scheme of mixed solid-phase solution-phase peptide synthesis shown with L2-15 peptide as an example.

### 2.2.3. Electrophoretic mobility shift assay (EMSA)

The *in vitro* evaluation of the Tat peptidomimetics by electrophoretic mobility shift assay (EMSA) was performed in cooperation with Prof. G. Varani from the Department of Chemistry at the University of Washington, and Dr. Zafiria Athanassiou and Amy Davidson.

The electrophoretic mobility shift assay is one of the most common methodologies for identification and characterisation of protein complexes with nucleic acids. The EMSA technique is based on the observation that protein:DNA/RNA complexes migrate more slowly than free DNA/RNA molecules when subjected to non-denaturing polyacrylamide or agarose gel electrophoresis.[95,96] Because the rate of DNA/RNA migration is shifted or retarded upon protein binding, the assay is also referred to as a gel shift or gel retardation assay. For visualization purposes, the DNA/RNA fragment is usually radiolabeled with $^{32}P$ by incorporating [$\alpha$-$^{32}P$]dNTPs during run-off transcription with T7 polymerase. After electrophoresis, the distribution of species containing nucleic acid is determined, usually by autoradiography of [$\alpha$-$^{32}P$]-labeled nucleic acid.

The specificity of the observed DNA/RNA binding reaction can be evaluated using competition assays in which an excess of unlabelled tRNA is added together with the labelled probe. Generally, non-specific competitors can be present in a large excess over the specific probe. Such a competitor, widely used in protein:RNA studies, is tRNA from yeast or *E. coli*.

The strength of the binding is usually described using the dissociation constant ($K_d$). A low dissociation constant implies a high binding affinity. This number can be determined quantitatively by analysis of mobility shift assays. The dissociation constant is determined from the equation $K_d$=[RNA][P]/[P-RNA], where [RNA] is the free concentration of RNA, [P] is the concentration of free protein, peptide or other ligand, and [P-RNA] is the concentration of the complex. In practise, reactions are performed with an excess of protein (ligand) over the RNA. Assuming that upon binding the concentration changes insignificantly, the [P] can be taken as the protein (ligand) concentration used. The value of [RNA]/[P-RNA] is determined from the fraction of the RNA present in the unshifted (unbound) versus shifted (bound) bands in the mobility shift assay. If there is 1:1 binding stoichiometry, then $K_d$ equals the

protein/peptide concentration, when 50% of the RNA is bound and present in the shifted band, so [RNA]=[P-RNA].

To test the binding affinity of Tat peptidomimetics to BIV RNA and HIV RNA, radiolabelled RNAs were prepared using [$\alpha$-$^{32}$P] CTP and T7 polymerase by *in vitro* run-off transcription. The titration experiments were performed with varying concentrations of the Tat peptidomimetics, a constant concentration of radiolabelled TAR RNA (2nM) and unlabelled competitor tRNA (10'000-fold excess). The wide-type linear BIV Tat (Ac-$^{65}$SGPRPRGTRGKGRRIRR$^{81}$-$NH_2$ (BIV Tat (65-81)) or ADP-1$^{72}$ (Ac-$^{37}$SFTTKALGISYGRKKRRQRRRPPQGSQTHQVSLSKQ$^{72}$-$NH_2$ (HIV-1 Tat(37-72)) peptides were used as positive controls in all binding studies. The $K_d$ determined by EMSA with a large excess of tRNA for BIV Tat (65-81) peptide is 50 nM and 200 nM for ADP-1.

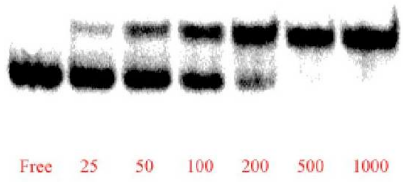



**Figure 19.** Example of a visualised gel in a EMSA of one of the Tat peptidomimetics (L2-15). Lower band corresponds to the free radiolabelled TAR RNA and upper band stands for the peptide:RNA complex. The concentrations of the peptide are shown in nanomolar. The $K_d$ of the interaction is estimated to be ~100 nM.

The L1 peptidomimetics were assayed by electrophoretic mobility shift assay (EMSA) primarily with their target BIV TAR RNA, but in addition with HIV-1 TAR to investigate the selectivity and to identify interesting inhibitors of the HIV-1 Tat-TAR interaction. The results are reported in the Table 5.

- **LIBRARY L2: Attempts to mutate the residues critical for binding, and to tune the β-turn region, and then to introduce multiple mutations based on the results obtained with single mutations.**

In the second round of optimisation, further attempts were made to mutate the residues that are involved in the critical contacts with TAR RNA, such as Arg1, Arg3 (Table 6). Furthermore, some amino acids that from the complex are not seen to be involved in binding, namely Val2, which is on the solvent exposed side, and Lys7 involved in β-turn formation but not in contact with RNA, were mutated. Additionally,

the Gly6 and Lys7 were exchanged to evaluate the conformational effects of residues involved in β-turn formation. Finally, the knowledge gained from single mutations was used to combine multiple mutations in order to evaluate if the affinity improvements are additive. As the Val12Ile in the first library improved the affinity 3-fold (150 vs. 50 nM) this mutation was introduced into all the L2 sequences.

**Table 6.** Library L2 of BIV Tat mimetics. Residues 1 and 12 are attached to the D-Pro-L-Pro template. $K_d$ values [μM] were determined by EMSA (nd=not determined).

| **Mimetic** | **Position** | | | | | | | | | | | | **$K_d$ BIV** | **$K_d$ HIV** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | | |
| **L2-01** | R | V | R | T | R | G | K | R | R | I | R | **I** | 0.1 | 0.5 |
| **L2-02** | **U** | V | R | T | R | G | K | R | R | I | R | **I** | >10 | nd |
| **L2-03** | **K** | V | R | T | R | G | K | R | R | I | R | **I** | >10 | 0.1 |
| **L2-04** | **O** | V | R | T | R | G | K | R | R | I | R | **I** | >10 | 0.5 |
| **L2-05** | **N** | V | R | T | R | G | K | R | R | I | R | **I** | **>**10 | >5 |
| **L2-06** | **Q** | V | R | T | R | G | K | R | R | I | R | **I** | 10 | >5 |
| **L2-07** | **Y** | V | R | T | R | G | K | R | R | I | R | **I** | 5 | 5 |
| **L2-08** | **L** | V | R | T | R | G | K | R | R | I | R | **I** | >5 | >5 |
| **L2-09** | **W** | V | R | T | R | G | K | R | R | I | R | **I** | >5 | >5 |
| **L2-10** | R | **Q** | R | T | R | G | K | R | R | I | R | **I** | 0.2 | 0.5 |
| **L2-11** | R | **Y** | R | T | R | G | K | R | R | I | R | **I** | 0.2 | 1 |
| **L2-12** | R | **W** | R | T | R | G | K | R | R | I | R | **I** | 0.5 | 5 |
| **L2-13** | R | **L** | R | T | R | G | K | R | R | I | R | **I** | 0.2 | 5 |
| **L2-14** | R | V | R | T | R | **N** | **G** | R | R | I | R | **I** | 0.5 | >5 |
| **L2-15** | R | V | R | T | R | **K** | **G** | R | R | I | R | **I** | 0.1 | 0.075 |
| **L2-16** | R | V | R | T | R | G | **R** | R | R | I | R | **I** | 0.2 | 1 |
| **L2-17** | R | V | R | T | R | G | **Y** | R | R | I | R | **I** | 0.2 | >5 |
| **L2-18** | R | V | R | T | R | G | **Q** | R | R | I | R | **I** | 0.2 | >5 |
| **L2-19** | R | V | R | T | R | G | **N** | R | R | I | R | **I** | 0.2 | >5 |
| **L2-20** | R | V | R | T | R | G | **O** | R | R | I | R | **I** | 0.1 | 5 |
| **L2-21** | R | V | R | T | R | G | **U** | R | R | I | R | **I** | >10 | >5 |
| **L2-22** | R | V | **Q** | T | R | G | K | R | R | I | R | **I** | >5 | >5 |
| **L2-23** | R | V | **N** | T | R | G | K | R | R | I | R | **I** | >5 | 5 |
| **L2-24** | R | **T** | R | T | R | G | K | R | R | I | R | **I** | 0.05 | 0.5 |
| **L2-25** | R | **N** | R | T | R | G | K | R | R | I | R | **I** | 0.1 | 0.25 |
| **L2-26** | R | V | R | T | R | G | K | R | R | I | **O** | **I** | 0.05 | 5 |
| **L2-27** | R | V | R | T | R | G | K | R | R | I | **K** | **I** | 0.2 | 5 |
| **L2-28** | R | V | R | T | R | G | K | R | **K** | I | **O** | **I** | 0.1 | 5 |
| **L2-29** | R | V | R | **Q** | R | G | K | R | R | I | R | **I** | 0.2 | >5 |
| **L2-30** | R | V | R | **V** | R | G | K | R | R | I | R | **I** | 0.2 | >5 |
| **L2-31** | R | V | R | **Y** | R | G | K | R | R | I | R | **I** | 0.2 | >5 |
| **L2-32** | R | **T** | R | **Q** | R | G | K | R | R | I | R | **I** | 0.1 | 5 |
| **L2-33** | R | **T** | R | T | R | G | K | R | R | I | **O** | **I** | 0.1 | 5 |
| **L2-34** | R | **T** | R | T | R | G | K | R | **K** | I | **O** | **I** | 0.1 | 5 |
| **L2-35** | R | **T** | R | **Q** | R | G | K | R | R | I | **O** | **I** | 0.1 | 1 |
| **L2-36** | R | **T** | R | **V** | R | G | K | R | R | I | **O** | **I** | 0.1 | 1 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **L2-37** | R | **T** | R | **Y** | R | G | K | R | R | I | **O** | **I** | 0.2 | 5 |
| **L2-38** | R | V | R | T | R | G | K | R | R | **F** | R | **I** | 1 | >5 |
| **L2-39** | R | **T** | R | T | R | G | K | R | R | **F** | R | **I** | 0.5 | >5 |
| **L2-40** | R | **T** | R | T | R | G | K | R | R | **F** | **O** | **I** | 0.2 | 5 |
| **L2-41** | R | V | R | **Q** | R | G | K | R | R | **F** | **O** | **I** | 0.5 | >5 |
| **L2-42** | R | **T** | R | **V** | R | G | K | R | **K** | **F** | **O** | **I** | 0.5 | >5 |
| **L2-43** | R | **T** | R | **Y** | R | G | K | R | **K** | **F** | **O** | **I** | 0.5 | >5 |

The peptidomimetic library 2 was prepared using a standard Fmoc solid phase peptide synthesis protocol on a MultiSyn Tech Syro II peptide synthesiser as described previously in paragraph 2.2.2. After purification by HPLC, the desired peptidomimetics were typically obtained in 20-40% overall yield, and each was >95% pure as determined by analytical HPLC and gave ESI-MS data consistent with the calculated masses (see Appendix 2 for analytical data).

The results of EMSA of peptides from both libraries established a structure-activity-relationship for the BIV Tat peptidomimetics, which will be discussed in detail in the Discussion Section (Paragraph 2.3.1) It is noteworthy that these two rounds of optimisation gave peptides with only about 3-fold increased affinities to BIV TAR RNA (peptides L1-02, L1-07, L1-15, L2-24, L2-26) in comparison to lead compound BIV2. This shows that the concept used in the design of BIV0-BIV8 peptides delivered already a very potent, almost optimal RNA ligand (BIV2). The extensive mutational studies gave results consistent with the structural information obtained for the BIV2-BIV TAR RNA complex. In general, only small changes in the character of amino acid residues in the BIV2 sequence were tolerated and the gain in the activity was not large.
To provide a measure of the intrinsic affinities of these mimetics for BIV TAR, without competition from lower-affinity non-specific binding sites, as provided by tRNA, a standard used in EMSA, the binding assays were repeated in the absence of tRNA. Although several peptidomimetics exhibited high affinity for TAR RNA even in the presence of a large excess of tRNA, the $K_d$'s are typically lower (tighter binding) by at least 1 order of magnitude in the absence of tRNA (Table 7). These data demonstrate that nanomolar ligands of BIV TAR have been discovered that are capable of competing with a much larger protein for binding to the TAR RNA, and bind to their target even in the presence of a 10'000-fold excess of non-cognate tRNA.

**Table 7.** $K_d$ values for selected mimetics binding to TAR RNA determined by EMSA in the presence and absence of tRNA.

| **Mimetic** | **$K_d$ [μM] BIV TAR EMSA with tRNA** | **$K_d$ [μM] BIV TAR EMSA without tRNA** |
|---|---|---|
| **BIV2** | 0.15 | 0.15 |
| **L2-01** | 0.1 | 0.07 |
| **L2-15** | 0.1 | 0.005 |
| **L2-24** | 0.05 | 0.03 |
| **L2-29** | 0.2 | 0.02 |
| **L2-26** | 0.05 | 0.005 |
| **L2-27** | 0.1 | 0.005 |
| **L2-28** | 0.1 | 0.01 |
| **L2-32** | 0.1 | 0.03 |
| **L2-33** | 0.05 | 0.02 |
| **L2-34** | 0.05 | 0.007 |
| **L2-35** | 0.05 | 0.007 |

### 2.2.4. HIV Tat peptidomimetics

By assaying the BIV Tat peptidomimetics with HIV-1 TAR RNA, several peptides, summarised in Table 8, that bind to the RNA with affinity comparable with ADP-1, were discovered. Importantly, peptides that bind the HIV and BIV TAR are distinct, and even small changes in peptide sequence lead to large losses in activity, proving the high specificity of the ligands.

**Table 8.** BIV Tat peptidomimetics with high affinity to HIV-1 TAR.

| **Mimetic** | **Position** | | | | | | | | | | | | **$K_d$ [nM] HIV-1 TAR** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | |
| L1-06 | R | V | R | T | R | G | K | R | R | I | R | **R** | 150 |
| L1-07 | R | **T** | R | T | R | G | K | R | R | I | R | V | 150 |
| L2-03 | **K** | V | R | T | R | G | K | R | R | I | R | **I** | 100 |
| L2-15 | R | V | R | T | R | **K** | **G** | R | R | I | R | **I** | 75 |

#### 2.2.4.1. 14-mer HIV-1 Tat peptidomimetics

- **LIBRARY L3: Further single mutations of the residues in BIV2 thought to be not critical for the binding to HIV TAR RNA.**

By assaying previously prepared BIV Tat peptidomimetics several peptides were discovered to be potent and specific ligands for HIV-1 TAR RNA (Table 5 and Table 6, summary in Table 8). Among them, L2-15 exhibits the highest affinity for HIV-1 TAR (75 nM), 3-fold higher than the peptide ADP-1 (200 nM), and this was used as a lead compound for further optimisation.

L2-15 was the only peptide in the whole series with a β-turn defined by Lys6-Gly7. This feature made it unique, suggesting that it is important for the interaction with HIV-1 TAR RNA. In an attempt to further investigate the turn region, the mutations Lys6Arg and Lys6hArg (peptides L3-04, L3-06) were studied (Table 9). By these changes the positive charge of the side chain was kept, but the guanidine group was introduced, and with two different lengths of side chain (Arg and hArg=homo-Arg). The previous mutation studies suggested that the positive charge in the turn region is important for the interaction with HIV-1, but not for BIV TAR RNA. Additionally, in two peptides, L3-03 and L3-05, a disulfide bridge was introduced at positions 2 and 11 to stabilize a β-hairpin conformation. The positions 2 and 11 in the

β-hairpin conformation correspond to a non-hydrogen bonding site, and secondly in the BIV2-TAR RNA complex, they are solvent exposed and do not contact the RNA. To keep the overall charge of these peptides constant, in addition to incorporation of two cysteines, the amino acid at position 10 was mutated (Ile10Arg and Ile10Lys, for L3-03 and L3-05, respectively). No structure is available of a ligand larger than arginamide in complex with HIV-1 TAR, so by the introduction of a disulfide bridge, which further stabilizes the β-hairpin conformation, we would evaluate if this conformation is important for the RNA recognition.

It is noteworthy, that for the first time a single mutation was introduced in the D-Pro-L-Pro template (D-Pro→D-Arg) in the peptide L3-01. In the structure of the BIV2-TAR RNA complex, the template makes no significant contacts to the BIV TAR RNA, in particular D-Pro protrudes out of the binding site into the solvent. Once again, having little structural information on the ligand-HIV-RNA complex, it was hypothesized that the template region may be in closer proximity to the HIV RNA than in the BIV system. By introducing the long, positively charged side chain of D-Arg, it was hoped to test this hypothesis. Although introducing a $^{D}$Pro13$^{D}$Arg mutation may have a slightly destabilizing effect on the hairpin conformation (more freedom in the ϕ angle for Arg in comparison to Pro), this effect might be minimized or even overcome if the guanidine group interacts with the RNA. Importantly, the stereochemistry of the mutated amino acid was kept as D.

**Table 9.** Library L3 of HIV-1 Tat mimetics. $K_d$ values determined by EMSA.

| **Mimetic** | **Position** | | | | | | | | | | | | **Template** | **$K_d$ [nM] HIV-1 TAR** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | | |
| **L3-01** | R | V | R | T | R | K | G | R | R | I | R | I | **D-Arg**-L-Pro | 100 |
| **L3-02** | R | V | R | T | R | K | G | R | R | **R** | R | I | D-Pro-L-Pro | 100 |
| **L3-03** | R | **C** | R | T | R | K | G | R | R | **R** | **C** | I | D-Pro-L-Pro | 300 |
| **L3-04** | R | V | R | T | R | **R** | G | R | R | I | R | I | D-Pro-L-Pro | 100 |
| **L3-05** | R | **C** | R | T | R | K | G | R | R | **K** | **C** | I | D-Pro-L-Pro | 75 |
| **L3-06** | R | V | R | T | R | **hR** | G | R | R | **K** | R | I | D-Pro-L-Pro | 100 |

The peptidomimetic library 3 was prepared using a standard Fmoc solid phase peptide synthesis protocol on a MultiSyn Tech Syro II peptide synthesiser, as described previously in paragraph 2.2.2.

In the case of peptides containing an intramolecular disulfide bridge, the oxidation was performed on the pre-purified cyclic peptides containing cysteine

residues protected with acetamidomethyl (Acm) protecting groups, in a mixture of AcOH:$H_2O$ (4:1) and 10 equivalents of iodine for 30 min. The quenching of the excess of iodine was achieved with 10 mM ascorbic acid in 50 mM citrate buffer (pH=5.0) (Figure 20).

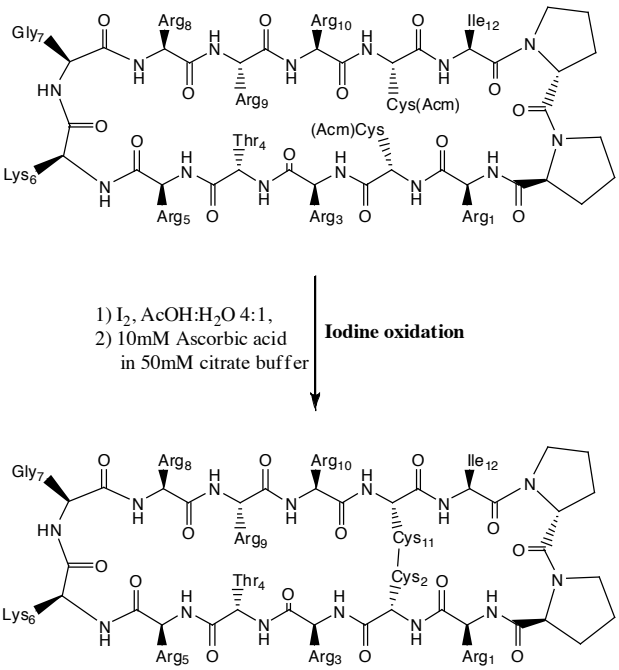



**Figure 20**. Iodine oxidation of the cysteine residues protected with an Acm group. The scheme is shown for L3-03 peptide as an example**.**

After purification by HPLC, the desired peptidomimetics were typically obtained in 20-40% overall yield, and each was >95% pure as determined by analytical HPLC and gave ESI-MS data consistent with the calculated masses (see Appendix 2 for analytical data).

Unfortunately, the small library did not afford a peptide with affinity higher than the lead peptide, L2-15 ($K_d$=75 nM). The mutations introduced proved not to be critical, all the peptides remain active. Mutations of Lys6 into Arg and hArg led to a slight decrease of affinity, showing that the electrostatic interaction between the positively charged residue and the RNA is probably more important than the specific hydrogen bonding ability that a guanidinium group could facilitate. Furthermore, neither the introduction of the disulfide bridge nor the replacement of D-Pro by D-Arg, had any significant effects on the potency of the peptides.

### 2.2.4.2. Extended HIV-1 Tat peptidomimetics

- **Attempt to introduce more contacts to the HIV-1 TAR RNA by extending the length of the peptidomimetic**

The first library of peptidomimetics proved that 14-amino acid cyclic peptides with stable β-hairpin conformations can be potent binders to HIV-1 TAR RNA. However, all attempts at producing ligands with higher affinity than the lead compound, L2-15, have so far failed

The mechanism of the Tat-TAR interaction in HIV-1, clearly shows that the cooperative binding of cyclin T1 to Tat protein and the TAR RNA loop region is necessary for effective transactivation. The cyclin T1-Tat heterodimer directly binds to the bulge region of TAR RNA, and this binding facilities the interactions of the cyclin T1-TAR RNA in the loop region.[76] It was hypothesised that by extending the peptide sequences from 14 to 18 residues the surface of the intermolecular contacts between the mimetic and the RNA might be increased. Furthermore, it was envisioned that the contacts of this extended loop with the tip of the HIV-1 TAR RNA might result in inhibition of the cyclin T1 and TAR RNA interaction. The first peptide designed in this fashion was based on the sequence of BIV2, as its structure was at this point of the project the best studied one and its interaction with the RNA well understood. As the extension of the loop would bring more flexibility to the peptide structure, apart from the D-Pro-L-Pro template, a disulfide cross-link at a non-hydrogen bonding position (4 and 13) was introduced to further stabilize a hairpin conformation (Figure 21).

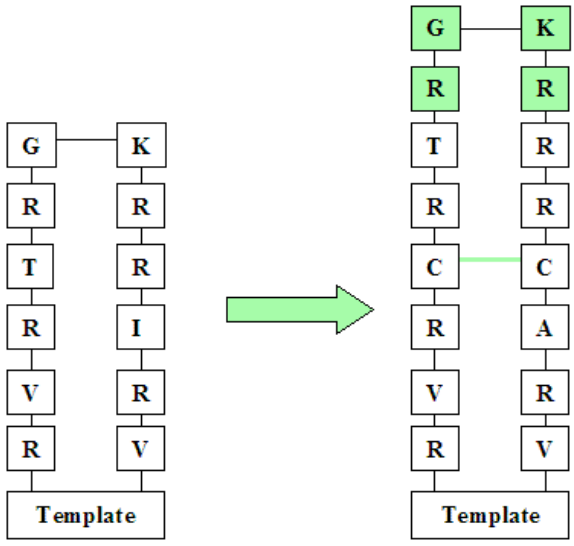



**Figure 21.** The first long loop peptide, called L4-01 (right), based on the BIV2 sequence (left).

The peptide L4-01 was prepared using a standard Fmoc solid phase peptide synthesis protocol on an ABI 433A peptide synthesiser, as described previously in paragraph 2.2.2. The formation of the disulfide cross-link was performed by oxidation on the pre-purified cyclic peptide containing cysteine residues protected with Acm groups, discussed in detail in paragraph 2.2.4.1. After purification by HPLC, the desired peptidomimetic was obtained in 38% yield. It was >95% pure as determined by analytical HPLC and gave ESI-MS data consistent with the calculated mass (see Appendix 2 for analytical data).

The EMSA performed with this peptide revealed that this longer loop peptide was as potent as ADP-1 ($K_d$=200 nM) in binding to HIV-1 TAR RNA. L4-01 was a first attempt to increase the contact surface of the ligand and RNA and its sequence was based on BIV2, which was not the best ligand for HIV-1 TAR RNA ($K_d$=500 nM) Therefore this finding was very encouraging for further studies.

- **LIBRARY L4: Optimisation of the lead longer-loop peptide**

On the basis of the encouraging result obtained with the longer-loop peptide, L4-01, a new set of 18-residue peptides were designed and assayed. The sequences were based on that of L2-15, so far the most potent HIV-1 TAR RNA ligand ($K_d$=75 nM). Multiple mutations were made to the amino acids at positions 6-11, including the β-turn defining residues (Table 10). The cysteines forming the disulfide bridge, necessary for the stabilization of the β-hairpin conformation, were incorporated at two different non-hydrogen bonding positions: **i)** in the first 7 peptides, L4-02—L4-08 at positions 4 and 13, the same as for L4-01, **ii)** in L4-09—L4-13 at positions 2 and 15.

**Table 10.** Library L4 of 18-mer HIV Tat peptidomimetics. Residues 1 and 16 are attached to the D-Pro-L-Pro template. The $K_d$ values were determined by standard EMSA.

| **Mimetic** | **Position** | | | | | | | | | | | | | | | | **$K_d$ [μM]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | ***13*** | **14** | ***15*** | **16** | |
| **L4-01** | **R** | **V** | **R** | ***C*** | **R** | **T** | **R** | **G** | **K** | **R** | **R** | **R** | ***C*** | **A** | **R** | **V** | **0.2** |
| **L4-02** | R | V | R | ***C*** | R | Q | R | K | G | R | A | R | ***C*** | I | R | I | 0.2 |
| **L4-03** | R | V | R | ***C*** | R | K | L | K | G | Q | T | R | ***C*** | I | R | I | >5 |
| **L4-04** | R | V | R | ***C*** | R | K | R | $^{D}$R | G | Q | T | R | ***C*** | I | R | I | 0.075 |
| **L4-05** | R | V | R | ***C*** | R | Q | R | G | P | G | K | R | ***C*** | I | R | I | >5 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **L4-06** | R | V | R | *C* | R | K | Q | G | P | G | T | R | *C* | I | R | I | >5 |
| **L4-07** | R | V | R | *C* | R | R | K | G | P | G | Q | R | *C* | I | R | I | 0.15 |
| **L4-08** | R | V | R | *C* | R | K | R | G | P | G | Q | R | *C* | I | R | I | >5 |
| **L4-09** | R | *C* | R | V | R | Q | R | K | G | R | A | R | K | I | *C* | I | 0.15 |
| **L4-10** | R | *C* | R | V | R | K | Q | G | P | G | T | R | K | I | *C* | I | >5 |
| **L4-11** | R | *C* | R | V | R | K | Q | G | P | G | K | R | K | I | *C* | I | >5 |
| **L4-12** | R | *C* | R | V | R | K | R | G | P | G | Q | R | K | I | *C* | I | >5 |
| **L4-13** | R | *C* | R | V | R | R | K | G | P | G | A | R | K | I | *C* | I | 0.2 |

The peptidomimetic library 4 was prepared using a standard Fmoc solid phase peptide synthesis protocol on a MultiSyn Tech Syro II peptide synthesiser, as described previously in paragraph 2.2.2. The intramolecular disulfide bridge, was formed by the oxidation of the pre-purified cyclic peptides containing cysteine residues protected with Acm groups, as previously described in detail in paragraph 2.2.4.1. After purification by HPLC, the desired peptidomimetics were typically obtained in 20-40% overall yield, and each was >95% pure as determined by analytical HPLC and gave ESI-MS data consistent with the calculated masses (see Appendix 2 for analytical data).

The L4 library containing 18-amino acid hairpins gave some ligands with higher affinity for TAR RNA than L4-01, namely L4-04, L4-07 and L4-09. As multiple mutations have been made and the sequences vary extensively, it is not easy to discuss the importance of individual amino acid residues. However, some useful conclusions can be drawn, which might aid future work. The position of the disulfide bridge appears not to be of importance, as in both subgroups potent peptides are identified. Various β-turn defining amino acids (KG, $^{D}$RG, GP) were tested and all of the combinations gave active peptides. L4-03 and L4-04 vary only at positions 7 and 8 (LK vs. R$^{D}$R) but this has a dramatic effect on the affinity (>5 μM vs. 75 nM). The only difference between L4-07 and L4-08 is the exchange of the amino acids at positions 6 and 7 (RK vs. KR) but again, the affinity diminishes drastically. Similar observations were made for L4-12 and L4-13, however, Lys6-Arg7 present in L4-04 are apparently tolerated well in this peptide.

In addition to the peptides containing 18 amino acid residues, it was decided to design several peptides with different loop lengths: **i**) 15-mers (L4-14—L4-19), and **ii**) 16-mers (L4-20—L4-25), summarised in Table 11 and Table 12. Varying the loop length (15-mer, 16-mer vs. 18-mer) would help to evaluate the importance of this extension, and the possibility of picking up additional contacts with the tip of the RNA loop.

**Table 11.** Library L4 of 15-mer HIV Tat peptidomimetics. Residues 1 and 13 are attached to the D-Pro-L-Pro template. The $K_d$ values were determined by standard EMSA.

| **Mimetic** | **Position** | | | | | | | | | | | | | **$K_d$ [μM]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | ***13*** | |
| **L4-14** | R | V | R | C | R | K | K | G | K | C | I | R | I | >5 |
| **L4-15** | R | V | R | C | R | R | R | G | R | C | I | R | I | >5 |
| **L4-16** | R | V | R | T | R | K | K | G | K | R | I | R | I | 1 |
| **L4-17** | R | V | R | C | R | K | K | $^{D}$P | K | C | I | R | I | >5 |
| **L4-18** | R | V | R | T | R | K | K | $^{D}$P | K | R | I | R | I | 0.75 |
| **L4-19** | R | V | R | V | R | K | K | G | R | T | I | R | I | >5 |

Analysis of binding of the 15-residue peptides to HIV TAR RNA by EMSA, showed that they do not bind, or bind weakly. Most likely the modified structure of the hairpin has significantly effected the hairpin stability and/or the interactions with the RNA.

**Table 12.** Library L4 of 16-mer HIV Tat peptidomimetics. Residues 1 and 14 are attached to the D-Pro-L-Pro template. The $K_d$ values were determined by standard EMSA.

| **Mimetic** | **Position** | | | | | | | | | | | | | | **$K_d$ [nM]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | ***13*** | **14** | |
| **L4-20** | R | V | R | C | R | K | R | K | G | R | C | I | R | I | 500 |
| **L4-21** | R | V | R | C | R | R | K | R | G | R | C | I | R | I | 300 |
| **L4-22** | R | V | R | T | R | Q | R | K | G | R | R | I | R | I | 200 |
| **L4-23** | R | V | R | V | R | K | R | K | G | R | Q | I | R | I | 250 |
| **L4-24** | R | V | R | Q | R | K | R | K | G | R | V | I | R | I | 200 |
| **L4-25** | R | V | R | C | R | K | R | K | $^{D}$P | R | C | I | R | I | 300 |

All the 16-amino acid peptides exhibit moderate binding to TAR RNA, including peptides with and without the disulfide bridge present. So apparently, any additional stabilization of the β-hairpin has no significance on the interaction with the RNA. At positions 4 and 11 various amino acids were incorporated and the minor changes in affinity proved that these changes are not critical for binding. The replacement of Gly9 by D-Pro (L4-25), also brings no significant improvement in affinity.

In summary, only peptides of 18-amino acid residues with the template, exhibit high potency among all the compounds tested and, therefore, they were chosen for further optimisation.

### 2.2.4.3. Structure-activity-relationship studies on lead compounds L4-04 and L2-15

- **Attempts to evaluate the importance of the template and disulfide bridge.**

L4-04 has been identified as the most potent peptide among the 18-amino acid ligands. Since there is no structure of the ligand/HIV-1 TAR RNA available, we performed systematic studies on this lead compound. In the first step, the importance of the template and disulfide bridge was evaluated. Three peptides were designed: **i**) L4-04a, stabilized by a disulfide bridge but without the template, **ii**) L4-04b, stabilized by the template, but without a disulfide cross-link, and finally **iii**) L4-04c, without any of the constraints, i.e. a linear peptide. In the peptides without disulfide bridge, the cysteine residues were replaced by aminobutyric acid (Abu) (Table 13).

These peptidomimetics were prepared using a standard Fmoc solid phase peptide synthesis protocol on an ABI 433A peptide synthesiser, as described previously in paragraph 2.2.2 and the intramolecular disulfide bridge was formed as described in paragraph 2.2.4.1. After purification by HPLC, the desired peptidomimetics were typically obtained in 20-40% overall yield, and each was >95% pure as determined by analytical HPLC and gave ESI-MS data consistent with the calculated masses (see Appendix 2 for analytical data).

**Table 13.** Derivatives of L4-04 with and/or without D-Pro-L-Pro template and disulfide bridge. The $K_d$ values were determined by standard EMSA. Abu=aminobutyric acid

| **Mimetic** | **Position** | | | | | | | | | | | | | | | **Template** | **$K_d$ [nM]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | ***13*** | **14** | ***15*** | **16** | |
| **L4-04** | **R** | **V** | **R** | ***C*** | **R** | **K** | **R** | **$^{D}$R** | **G** | **Q** | **T** | **R** | ***C*** | **I** | **R** | **I** | **D-Pro-L-Pro** | **75** |
| **L4-04a** | **Ac-R** | V | R | *C* | R | K | R | $^{D}$R | G | Q | T | R | *C* | I | R | **I-$NH_2$** | -- | 200 |
| **L4-04b** | R | V | R | **Abu** | R | K | R | $^{D}$R | G | Q | T | R | **Abu** | I | R | I | D-Pro-L-Pro | 100 |
| **L4-04c** | **Ac-R** | V | R | **Abu** | R | K | R | $^{D}$R | G | Q | T | R | **Abu** | I | R | **I-$NH_2$** | -- | 200 |

The peptides not stabilized by the template exhibit a reduced potency by more than 2-fold, which suggests that the stabilization by D-Pro-L-Pro is an important factor. The absence of the second constraint (the disulfide bridge) has a much smaller effect on the affinity; the $K_d$ is only slightly worse in comparison to L4-04.

A similar comparison was performed for L2-15, the most potent of the 14-mer peptides, where the linear peptide without template was synthesised and assayed (Table 14).

**Table 14.** Comparison between L2-15 and its derivate without the template, L2-15a. The $K_d$ values were determined by standard EMSA.

| Mimetic | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | Template | $K_d$ [nM] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **L2-15** | **R** | **V** | **R** | **T** | **R** | **K** | **G** | **R** | **R** | **I** | **R** | **I** | **D-Pro-L-Pro** | 75 |
| **L2-15a** | **Ac-R** | V | R | T | R | K | G | R | R | I | R | **I-$NH_2$** | -- | 75 |

The linear peptide exhibits exactly the same affinity towards TAR RNA as the parent mimetic L2-15. This result suggests that the template in 14-mer peptides has no significance in contrast to the situation in the longer-loop peptides discussed above. However, the cyclic peptides are much more stable against proteolysis (Paragraph 2.2.8.1) so it is important to keep the template, when considering other properties.

- **Attempts to evaluate the importance of arginine residues in L4-04 by single mutation to lysine.**

The single mutations of arginine residues to lysine residues in the lead compound, L4-04, was performed to evaluate the significance of the guanidinium groups, which can form specific hydrogen bonding with base pairs and/or the RNA backbone.

**Table 15.** The HIV-1 Tat peptidomimetics of the L4-04K library. Single Arg→Lys mutations were introduced into the L4-04 lead compound. Residues 1 and 16 are attached to the D-Pro-L-Pro template. The $K_d$ values were determined by standard EMSA.

| Mimetic | Position | | | | | | | | | | | | | | | | $K_d$ [nM] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *1* | 2 | *3* | 4 | *5* | 6 | *7* | 8 | *9* | 10 | *11* | 12 | *13* | 14 | *15* | 16 | |
| **L4-04** | **R** | **V** | **R** | **C** | **R** | **K** | **R** | **$^{D}$R** | **G** | **Q** | **T** | **R** | **C** | **I** | **R** | **I** | **75** |
| **L4-04K01** | **K** | V | R | C | R | K | R | $^{D}$R | G | Q | T | R | C | I | R | I | 100 |
| **L4-04K02** | R | V | **K** | C | R | K | R | $^{D}$R | G | Q | T | R | C | I | R | I | 50 |
| **L4-04K03** | R | V | R | C | **K** | K | R | $^{D}$R | G | Q | T | R | C | I | R | I | 200 |
| **L4-04K04** | R | V | R | C | R | K | **K** | $^{D}$R | G | Q | T | R | C | I | R | I | 200 |
| **L4-04K05** | R | V | R | C | R | K | R | $^{D}$R | G | Q | T | **K** | C | I | R | I | 75 |
| **L4-04K06** | R | V | R | C | R | K | R | $^{D}$R | G | Q | T | R | C | I | **K** | I | >1000 |
| **L4-04K07** | R | V | R | C | R | K | R | **$^{D}$K** | G | Q | T | R | C | I | R | I | 75 |

These peptides were synthesised, purified and characterised using the methods described above, for the analytical data see Appendix 2.

The mutation Arg3Lys improves the affinity of the peptide, the positive charge of the lysine may produce a better fit in the complex than the guanidine group. The mutations Arg12Lys and $^{D}$Arg8$^{D}$Lys are perfectly tolerated and do not bring any improvement in affinity. Lysine replacements at positions 1, 5 and 7 seem to be well tolerated and bring only a small decrease in affinity. At these positions the guanidinium group is slightly preferred over the amine group. A dramatic drop in affinity is observed by the mutation Arg15Lys, where the affinity is completely lost. This result indicates that arginine with its guanidinium group at this position is critical for effective binding to the TAR RNA.

- **Attempts to evaluate the importance of the length of the arginine side chain in L4-04 and L2-15 by single mutations Arg→hArg.**

In addition to the Arg→Lys mutations described above, the single arginine residues were replaced stepwise by homoarginine (hArg) to evaluate the significance of the length of the side chain bearing guanidinium group. It was hypothesized that at some positions extending the side chain might bring an increase of affinity due to better contact with base pairs or the RNA backbone.

**Table 16.** Single Arg→hArg mutations were introduced into the L4-04 lead compound. Residues 1 and 16 are attached to the D-Pro-L-Pro template. The $K_d$ values were determined by standard EMSA. hR=homoarginine.

| **Mimetic** | **Position** | | | | | | | | | | | | | | | | **$K_d$ [nM]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | ***13*** | **14** | ***15*** | **16** | |
| **L4-04** | **R** | **V** | **R** | ***C*** | **R** | **K** | **R** | **$^{D}$R** | **G** | **Q** | **T** | **R** | ***C*** | **I** | **R** | **I** | **75** |
| **L4-04hR01** | **hR** | V | R | *C* | R | K | R | $^{D}$R | G | Q | T | R | *C* | I | R | I | 350 |
| **L4-04hR02** | R | V | **hR** | *C* | R | K | R | $^{D}$R | G | Q | T | R | *C* | I | R | I | 150 |
| **L4-04hR03** | R | V | R | *C* | **hR** | K | R | $^{D}$R | G | Q | T | R | *C* | I | R | I | 300 |
| **L4-04hR04** | R | V | R | *C* | R | K | **hR** | $^{D}$R | G | Q | T | R | *C* | I | R | I | 100 |
| **L4-04hR05** | R | V | R | *C* | R | K | R | $^{D}$R | G | Q | T | **hR** | *C* | I | R | I | 100 |
| **L4-04hR06** | **R** | V | R | *C* | R | K | R | $^{D}$R | G | Q | T | R | *C* | I | **hR** | I | 150 |

*Synthesis of Fmoc-N,N'-bis-Boc-homoarginine (**20**).* The Fmoc/Boc protected homoarginine (**20**) was prepared as shown in Figure 22. The Boc protection of commercially available guanidine hydrochloride to form *N,N'*-bis-Boc-guanidine (**18**)

was performed according to the procedure developed by Feichtinger.[97] Introduction of the triflate group gave *N,N'*-bis-Boc-*N''*-Tf-guanidine (**19**). This guanidinylating reagent was used in a reaction with the primary amine of Fmoc-protected lysine following the slightly modified description by Tamaki.[98] The overall yield was 40%.

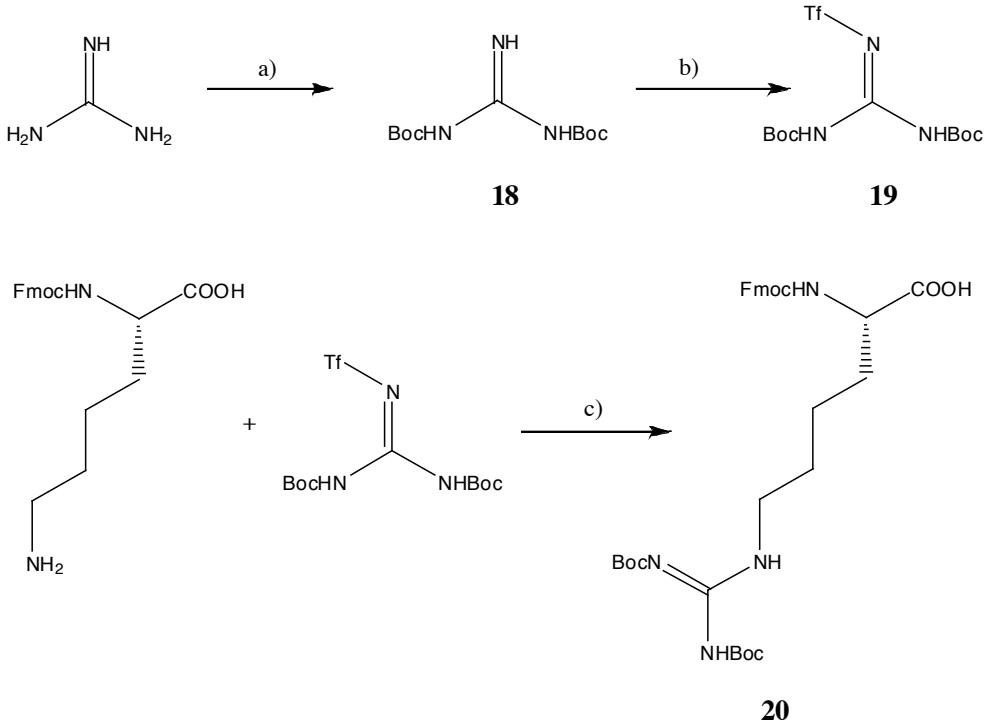



**Figure 22.** Scheme of synthesis of Fmoc-*N,N'*-bis-Boc-homoarginine. Reagents a) $Boc_2O$, NaOH, 1,4-dioxane; 66%, b) $Tf_2O$, DIEA, DCM; 65%, c) TEA, DCM; 95%.

The peptidomimetics with hArg incorporated were prepared using a standard Fmoc solid phase peptide synthesis protocol on a MultiSyn Tech Syro II peptide synthesiser, as described previously in paragraph 2.2.2. The coupling of the hArg was performed with a reduced amount of this amino acid (2 eq) and its completeness was monitored by the Kaiser test. The intramolecular disulfide bridge was formed by the oxidation of the pre-purified cyclic peptides containing cysteine residues protected with Acm groups, as previously described in detail in paragraph 2.2.4.1. After purification by HPLC, the desired peptidomimetics were typically obtained in 10-25% overall yield, and each was >95% pure as determined by analytical HPLC and gave ESI-MS data consistent with the calculated masses (see Appendix 2 for analytical data).

In all the cases the peptides bearing mutations are slightly weaker binders to TAR RNA than the lead compound, L4-04. The decrease in affinity is pronounced for positions Arg1 and Arg5 (more than 3-fold). Peptides carrying mutations at positions

Arg7, Arg12 and Arg15 maintain good affinity. Probably these arginine residues, as in BIV2, are on the solvent exposed face of the peptide and therefore are more tolerant to the changes.

The same approach was taken for the peptide L2-15. All the Arg→hArg mutations delivered peptides with slightly lower affinity compared to the parent peptide. Again, the arginine residues at positions 1, 3 and 5 are more affected by extension of the side chain, while the peptides with mutations of residues on the other side of the hairpin (Arg8, Arg9 and Arg11) retain the affinity (Table 17).

**Table 17.** Single Arg→hArg mutations were introduced into the L2-15 lead compound. Residues 1 and 12 are attached to the D-Pro-L-Pro template. The $K_d$ values were determined by standard EMSA. hR=homoarginine.

| **Mimetic** | **Position** | | | | | | | | | | | | **$K_d$ [nM]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | |
| **L2-15** | **R** | **V** | **R** | **T** | **R** | **K** | **G** | **R** | **R** | **I** | **R** | **I** | **75** |
| **L2-15hR01** | **hR** | V | R | T | R | K | G | R | R | I | R | I | 150 |
| **L2-15hR02** | R | V | **hR** | T | R | K | G | R | R | I | R | I | 200 |
| **L2-15hR03** | R | V | R | T | **hR** | K | G | R | R | I | R | I | 200 |
| **L2-15hR04** | R | V | R | T | R | K | G | **hR** | R | I | R | I | 75 |
| **L2-15hR05** | R | V | R | T | R | K | G | R | **hR** | I | R | I | 100 |
| **L2-15hR06** | R | V | R | T | R | K | G | R | R | I | **hR** | I | 100 |

The mutations of arginine to homoarginine introduced into both lead compounds, L2-15 and L4-04, show that extension of the side chain with an additional $-CH_2-$ group, does not facilitate any stronger contacts between guanidinium groups and TAR RNA. There are no significant changes in affinity, however, there is a trend noticeable - the mutations of the arginine residues 1, 3 and 5 are more affected by the extension, while other arginine residues are more tolerant with only a slight, if any, decrease in affinity. On the bases of these results we can suggest that the interaction between these peptides and HIV-1 TAR RNA is similar to that of BIV2-BIV TAR RNA and one side of the peptide predominantly contacts the RNA while the other side is rather solvent exposed.

- **LIBRARY L5: Further attempts to optimise the lead sequence of L4-04.**

From part of the L4 library containing 18-residue peptides, the lead compound, L4-04, has been identified. Attempts to evaluate the importance of intramolecular

constraints (the template and disulfide bridge) revealed that the template is more important for the binding affinity than the disulfide bridge, however, the linear peptide, L4-04c, still exhibits affinity to the RNA (Table 13).

Studying the importance of arginine residues present in L4-04 has shown so far that most of the changes made are well tolerated, with the exception of Arg15Lys, which causes a complete loss in affinity. The length of the arginine side chain has no major significance on the strength of the interaction between the peptide and the RNA. Nonetheless, structural information on the complex is still lacking, so only careful stepwise optimisation will enable us to gather more insights into this ligand-RNA recognition and allow us to produce peptides with higher affinity for HIV-1 TAR RNA. One priority was to investigate the importance of the turn region. In L4-04, D-Arg in the β-turn seems to play an important role. To evaluate its importance the mutations $^{D}$Arg8Arg and $^{D}$Arg8Lys were introduced. The combination of Lys-Gly has been found in the 14-residue peptidomimetics to be favourable over any other combination in the interaction with HIV-1 TAR. Gly-Pro is also known to stabilize the β-hairpin conformation, so it was considered worthwhile to introduce these changes into L4-04.

Peptide L4-04b showed that the disulfide bridge is not critical for the binding affinity (Table 13). It was therefore decided to replace the cysteine residues by threonine and alanine (L5-05), or to shift the disulfide to other non-hydrogen bonding positions (L5-06, L5-08 and L5-09). Additionally, multiple mutations were introduced at positions 6, 10, 11 and 15 combining the knowledge obtained from L4-01—L4-13 peptides. Finally, it was decided to transplant the turn region of L2-15 into the "longer-loop" peptides (L5-10—L5-12), to evaluate the effect on binding to HIV-1 TAR RNA.

The peptidomimetic library L5 was prepared using a standard Fmoc solid phase peptide synthesis protocol on a MultiSyn Tech Syro II peptide synthesiser, as described previously in paragraph 2.2.2. The intramolecular disulfide bridge, present in the majority of the peptides, was formed by the oxidation of the pre-purified cyclic peptides containing cysteine residues protected with Acm groups, as previously described in detail in paragraph 2.2.4.1. After purification by HPLC, the desired peptidomimetics were typically obtained in 20-40% overall yield, and each was >95% pure as determined by analytical HPLC and gave ESI-MS data consistent with the calculated masses (see Appendix 2 for analytical data).

**Table 18.** The HIV-1 Tat peptidomimetics of the L5 library. Residues 1 and 16 are attached to the D-Pro-L-Pro template. The $K_d$ values were determined by standard EMSA.

| **Mimetic** | **Position** | | | | | | | | | | | | | | | | **$K_d$ [nM]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | ***13*** | **14** | ***15*** | **16** | |
| **L4-04** | **R** | **V** | **R** | ***C*** | **R** | **K** | **R** | **$^{D}$R** | **G** | **Q** | **T** | **R** | ***C*** | **I** | **R** | **I** | **75** |
| **L5-01** | R | V | R | *C* | R | K | R | R | G | Q | T | R | *C* | I | R | I | 100 |
| **L5-02** | R | V | K | *C* | R | K | R | K | G | Q | T | R | *C* | I | R | I | 100 |
| **L5-03** | R | V | R | *C* | R | K | L | $^{D}$R | G | Q | T | R | *C* | I | R | I | 250 |
| **L5-04** | R | V | R | *C* | R | R | K | $^{D}$R | G | Q | T | R | *C* | I | R | I | 100 |
| **L5-05** | R | V | R | T | R | K | R | $^{D}$R | G | Q | T | R | A | I | K | I | 150 |
| **L5-06** | R | *C* | R | V | R | K | R | $^{D}$R | G | Q | T | R | R | I | *C* | I | 200 |
| **L5-07** | R | V | R | *C* | R | K | R | G | P | G | A | R | *C* | I | K | I | >1000 |
| **L5-08** | R | *C* | R | V | R | R | K | K | G | R | A | R | K | I | *C* | I | 200 |
| **L5-09** | R | *C* | R | V | R | Q | R | G | P | G | A | R | K | I | *C* | I | >1000 |
| **L5-10** | R | V | R | *C* | R | T | R | K | G | R | A | R | *C* | I | K | I | 75 |
| **L5-11** | R | V | R | *C* | R | Q | R | K | G | R | R | A | *C* | I | R | I | 25 |
| **L5-12** | R | V | R | *C* | R | Q | R | K | G | R | R | I | *C* | I | R | I | 75 |

The results of EMSA with the L5-01—L5-12 peptides, shown in Table 18, gives more information about the importance of several amino acid residues, and in combination with the results obtained with peptides L4-01—L4-13, allows new conclusions to be drawn.

The comparison of the peptides with and without a disulfide bridge (L4-04, L4-04b and L5-05) confirms that the additional stabilization of the conformation of the 18-mer peptides leads to only minor changes in the binding affinity. The peptide without the bridge shows only a small decrease in affinity (2-fold). The position of the disulfide bridge is not a critical factor, but from all the peptides, the ones with the cysteines at positions 2 and 15 are generally 2 to 3-fold weaker binders to TAR RNA.

The sequence defining the tip of the loop has been modified in several peptides without much impact on the affinity. The D-Arg at position 8 may be exchanged for L-Arg and L-Lys, so neither the stereochemistry of the chiral centre, nor the guanidinium group are important at this position for the binding. Introduction of the Lys8-Gly9 at the turn (known from L2-15) results in retention or even an increase of potency (peptides L5-10, L5-11, L5-12). A β-turn defined by $-^{8}$GPG$^{10}$- works well with L4-07 and L4-13 (affinity of 150 and 200 nM, respectively), but in combination of the other mutations, brings a complete loss in affinity (L5-07 and L5-09).

Position 6 can accommodate different amino acid residues, including charged ones (lysine and arginine), and uncharged polar ones (glutamine or threonine). The

charge seems not to play a critical role; the peptides with Gln6 or Thr6 exhibit the same or even improved $K_d$'s.

Similar behaviour is observed at position 12, which in L5-11 and L5-12 is mutated into alanine or isoleucine, respectively, and the affinity is retained or improved. Interestingly, the arginine residue at position 15, which from the lysine-scanning library (Table 15) seems to be critical for binding and not exchangeable for lysine, in this library shows more tolerance towards this mutation (peptides L5-05 and L5-10). However, these peptides have other mutations, that may allow such a change and compensate for it by other contacts. It may be specially true for L5-05, which without the disulfide bridge exhibits more intrinsic flexibility than the constrained peptide, L4-04.

Two of the peptides that exhibited up-to-date the highest affinity towards HIV-1 TAR RNA in standard EMSA conditions with a large excess of tRNA, were assayed in EMSA in the absence of the tRNA (Table 19). The peptides exhibit affinities in the subnanomolar range and they are around 50-fold stronger in the absence of the unspecific competitor. These results show, that although we managed to design specific ligands of high affinity for TAR RNA, their partial binding to tRNA is unavoidable due to their cationic nature. However, it is noteworthy, that the binding affinities in the presence of 10'000-fold excess of competitor (tRNA) are still in the low nanomolar range, which makes them attractive leads for future development.

**Table 19.** $K_d$ values for selected mimetics binding to HIV-1 TAR RNA determined by EMSA in the presence and absence of tRNA.

| **Mimetic** | **$K_d$ [nM]** HIV-1 TAR EMSA **with tRNA** | **$K_d$ [nM]** HIV-1 TAR EMSA **without tRNA** |
|---|---|---|
| **L4-04K02** | 50 | 1 |
| **L5-11** | 25 | 0.5 |

#### 2.2.4.4. Electrophoretic mobility shift inhibition assay

In addition to the standard EMSA, the selected most potent HIV Tat peptidomimetics were assayed in competition with ADP-1, a native HIV-1 Tat-derived peptide, and a cyclin T1 bifunctional construct with Tat.[99] The aim of this experiment was to prove that the peptidomimetics have an ability to inhibit: **i**) the interaction between the RNA-binding domain of Tat (ADP-1) and its cognate RNA, and **ii**) the

interaction between the chimera of human cyclin T1 (hCycT1) and Tat and HIV-1 TAR RNA.

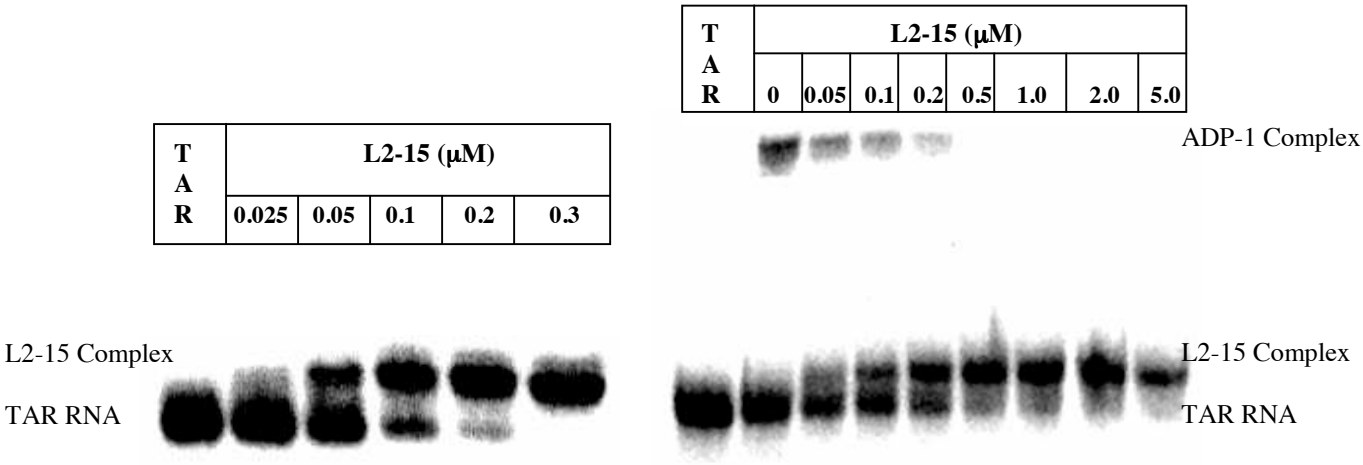



**Figure 23.** EMSA inhibition assay performed for L2-15. Left: Binding of L2-15 to HIV-1 TAR RNA (1 nM), right: Inhibition of the complex formation between the ADP-1 (150 nM) peptide and HIV-1 TAR RNA (1 nM) by L2-15.

The peptides exhibit a high potency to inhibit the interaction between ADP-1 and the hCytT1-Tat construct and its cognate TAR RNA (Figure 23and Table 20). That fact has a fundamental importance when considering the inhibition of viral transcription *in vivo*. Peptidomimetics are able to disturb not only the complex formation between HIV TAR RNA and ADP-1 but also the interaction of hCycT1 and Tat protein with the RNA. Since hCycT1 and Tat form a high-affinity complex with TAR RNA that results in the recruitment of Cdk9, formation of the positive transcription elongation factor-b complexes and reactivation RNA polymerase II, inhibition of this ternary complex formation is crucial for the inactivation of virus transcription.

**Table 20.** Summary of EMSA experiments, the standard direct binding assay, inhibition assay with ADP-1, and the inhibition assay with a chimera of hCycT1 and Tat.[99]

| Mimetic | $K_d$ [nM] | $IC_{50}$ [nM] ADP-1 | $IC_{50}$ [nM] Chimeric |
|---|---|---|---|
| **L2-15** | 75 | 100 | 250 |
| **L2-15a** | 75 | 100 | 100 |
| **L4-02** | 200 | 250 | 200 |
| **L4-04** | 75 | 100 | 100 |
| **L4-07** | 150 | 100 | 100 |
| **L4-09** | 150 | 150 | 150 |
| **L4-13** | 200 | 200 | 300 |

### 2.2.5. NMR analysis of selected HIV Tat peptidomimetics

The conformation of BIV2 was shown earlier to be a stable β-hairpin in aqueous solution.[54,93] To verify, that the other highly active 14-residue HIV-1 Tat peptidomimetics retained the same structure, NMR was used to characterise their structure in the absence of RNA (peptides under study listed in Table 21).

**Table 21**. Sequences of the selected 14-amino acid peptidomimetics, whose conformations have been studied in detail by NMR methods. Residues 1 and 12 are attached to the D-Pro-L-Pro template.

| **Mimetic** | **Position** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** |
| **BIV2** | **R** | **V** | **R** | **T** | **R** | **G** | **K** | **R** | **R** | **I** | **R** | **V** |
| L1-07 | R | **T** | R | T | R | G | K | R | R | I | R | V |
| L1-43 | R | V | R | T | R | G | K | R | R | I | **O** | V |
| L2-15 | R | V | R | T | R | **K** | **G** | R | R | I | R | **I** |

All $^1$H-NMR spectra of selected peptides were measured in aqueous solution (10%$D_2O$/$H_2O$, pH=2.3) if not otherwise indicated. The strategy of proton assignment relied on the identification of spin systems using COSY and TOCSY, followed by analysis of NOESY spectra to establish the sequential connectivities. Furthermore, H/D exchange experiments and a temperature series of $^1$H and TOCSY experiments were performed to investigate intramolecular hydrogen bonding and the conformational stability of the structure. Calculations of peptide structures were performed by using NOE-derived upper-distance limits and $J_{NH\text{-}H\alpha}$ coupling constants as restraints for a conformational search using DYANA.[100] MOLMOL was used for structural analysis and visualisation of the molecular structures.[101] These calculations were performed in collaboration with Dr. K. Moehle.

The differences between CαH chemical shift values determined experimentally and the random coil values are widely used as an indicator of secondary structure formation. In regular secondary structures of proteins, an up-field shift of amide NHs and CαHs are observed upon helix formation, and a down-field shift upon β-sheet formation.[102] Down-field shifts of Hα >0.1 ppm (Δδ), indicative of an extended β-sheet conformation, are observed for residues 1-5 and 8-12 in all the peptides studied (Figure 24).

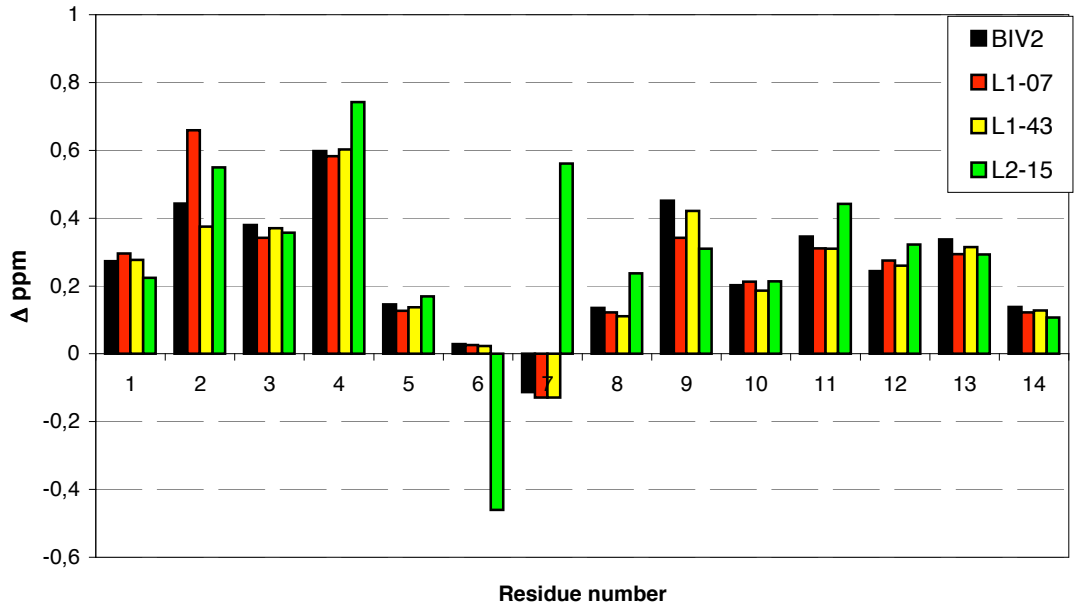



**Figure 24.** Hα chemical shifts relative to random coil chemical shifts of residues in peptidomimetics BIV2 (black), L1-07 (red), L1-43 (yellow) and L2-15 (green).

Peptides L1-07 and L1-43 (Table 21) vary from BIV2 only in one amino acid, but in both the β-turn is defined by Gly6-Lys7. L2-15 contains three mutations to the parent BIV2 sequence (Table 21). Especially the β-turn sequence is modified, which might significantly affect the hairpin stability and therefore this peptide has been studied in detail. It is well-known that the turn conformation plays a key role in the formation and stability of hairpin conformations in linear peptides. A comparison of $^{1}$H-spectra in the linear L2-15a (Table 14) and cyclic L2-15 peptides shows a large difference in the amide NH region. The spread of the amide proton signals for cyclic L2-15 and linear L2-15a (1.85 vs 0.43 ppm) indicates clearly the highly organised structure of the cyclic peptide in contrast to that in the unstructured linear peptide (Figure 25).

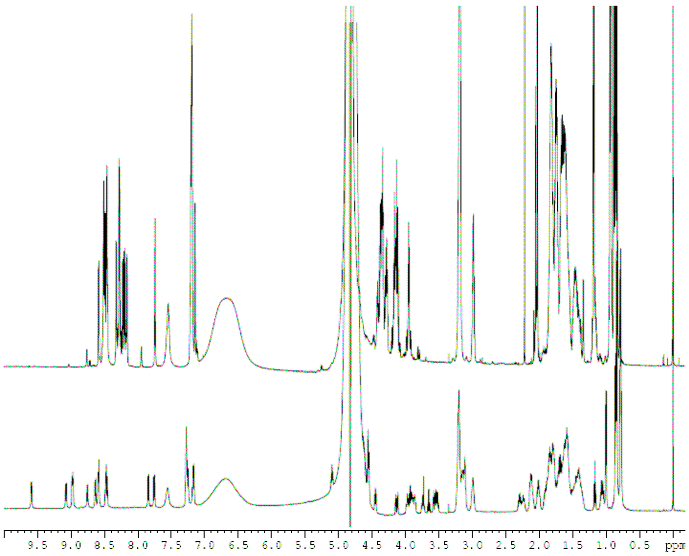



**Figure 25.** Comparison of the linear L2-15a peptide (without the template) [upper spectra] and the cyclic L2-15 [lower spectra].

The structure for L2-15 in aqueous solution was determined using 128 NOE upper distance limits. The ensemble of 20 structures had a backbone rmsd from the average structure of 0.75±0.36 Å (Table 22) indicating a highly structured molecule. Even prior to a formal structure calculation, conclusive evidence for a β-hairpin structure comes from the observation of long-range NOEs between residues on opposite strands of the hairpin, especially Hα-Hα NOEs between non-hydrogen bonding residue pairs and HN-HN NOEs between hydrogen bonding residue pairs on opposite strands. In L2-15, Hα-Hα NOEs are observed between Val2 and Arg11 and between Thr4 and Arg9; three HN-HN NOEs involve Arg1-Ile12, Arg3-Ile10, and Arg5-Arg8 connections. Furthermore, cross-strand Hα-HN NOEs were seen between Thr4 and Ile10 and between Val2 and Ile12. The average structures (Figure 26) confirm the predicted stable β-hairpin conformation in which the two strands of the antiparallel β-sheet are connected by a type I' turn between Lys6 and Gly7, with Lys6 occupying the left-hand region of conformational space.

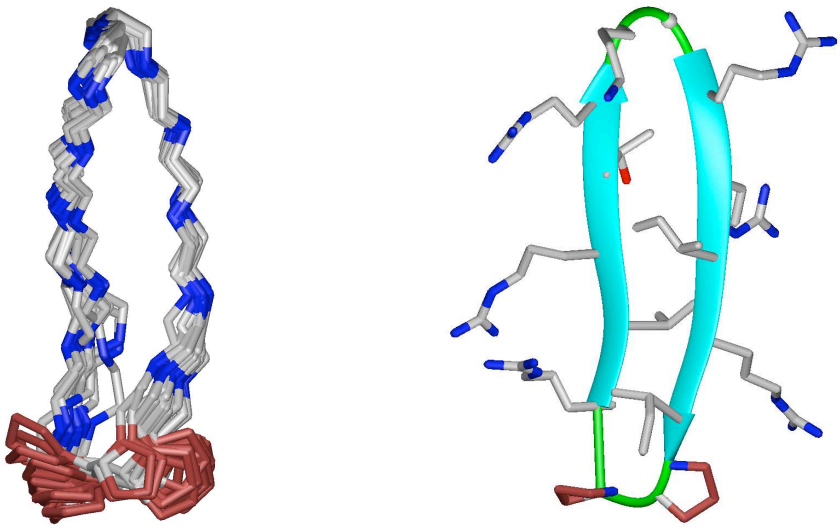


**Figure 26.** Left*:* Backbone superposition and representation of the final 20 NMR structures for **L2-15** with the template residues in red. Right: A typical NMR structure with the ribbon representation of the two strands.

A semiquantitative estimate of the population of folded structures can be obtained from the intensities of the cross-strand Hα-Hα NOEs, relative to the intensity of the Hα-Hα NOE of Gly7, as a fixed reference distance.[103] However, only the Thr4-Arg9 NOE could be clearly resolved in the $D_2O$ spectra, because the Val2-Arg11 NOE is in the proximity of the diagonal. With reference to the Hα-Hα separation of 2.32 Å found in regular antiparallel β-sheets in proteins (100% folded hairpin), the estimated hairpin population in L2-15 is then ~90%. Another method of estimating folded β-structures makes use of the $\alpha N(i,i+1)/\alpha N(i,I)$ NOE intensity ratio.[103] Values of ~2.3 correlate with random coil conformations, whereas larger values of ~8 are predicted for an ideal antiparallel β-sheet. The intensity ratios for residues in the predicted β-strands of L2-15 (1-5 and 8-11) are in the range of 4.4-11.4 and thus provide further support for highly populated β-conformations. The intensity ratio of <1 for the same pair of NOEs observed for Lys6 and Gly7 is consistent with the left-handed ($\alpha_L$) population in a type I' turn. The predominant occurrence of the type I' β-turn, with the atypical left-handed helical orientation of Lys6, is confirmed by characteristic NOE connectivities, in particular, by the sequential HN-HN (Lys6-Gly7 and Gly7-Arg8) and Hα-HN (Lys6-Arg8) NOEs, as well as the strong intramolecular Hα-HN NOEs of Lys6 and Gly7.

Long-range NOEs consistent with the proximity of the two antiparallel β-strands were also observed for other studied peptides L1-07, L1-43 and L2-15. They all

exhibit average β-hairpin structures in solution (Figure 27) that are very similar to those deduced earlier for BIV2 peptide (Figure 17). The results of the structural calculations of these peptides are summarised in Table 22 (see Appendix 6 for $^{1}$H-NMR assignments).

| **Mimetic** | **L1-07** | **L1-43** | **L2-15** |
|---|---|---|---|
| **NOE upper-distance limits** | 84 | 102 | 138 |
| Intraresidue | 31 | 37 | 56 |
| Sequential | 39 | 43 | 59 |
| Medium and Long-Range | 14 | 22 | 23 |
| **Residual Target Function Value (Å$^2$)** | 0.78±0.30 | 0.48±0.009 | 0.44±0.05 |
| **Mean rmsd values (Å)** | | | |
| All backbone atoms | 1.39±0.68 | 0.87±0.40 | 0.75±0.36 |
| All heavy atoms | 3.49±1.00 | 2.54±0.71 | 2.50±0.65 |
| **Residual NOE violations** | | | |
| Violation > 0.2 Å | 13 | 6 | none |
| Maximum violation (Å) | 0.44 | 0.24 | none |

**Table 22.** Experimental distance restraints and statistics for the final 20 NMR structures calculated for, L1-07, L1-43 and L2-15.

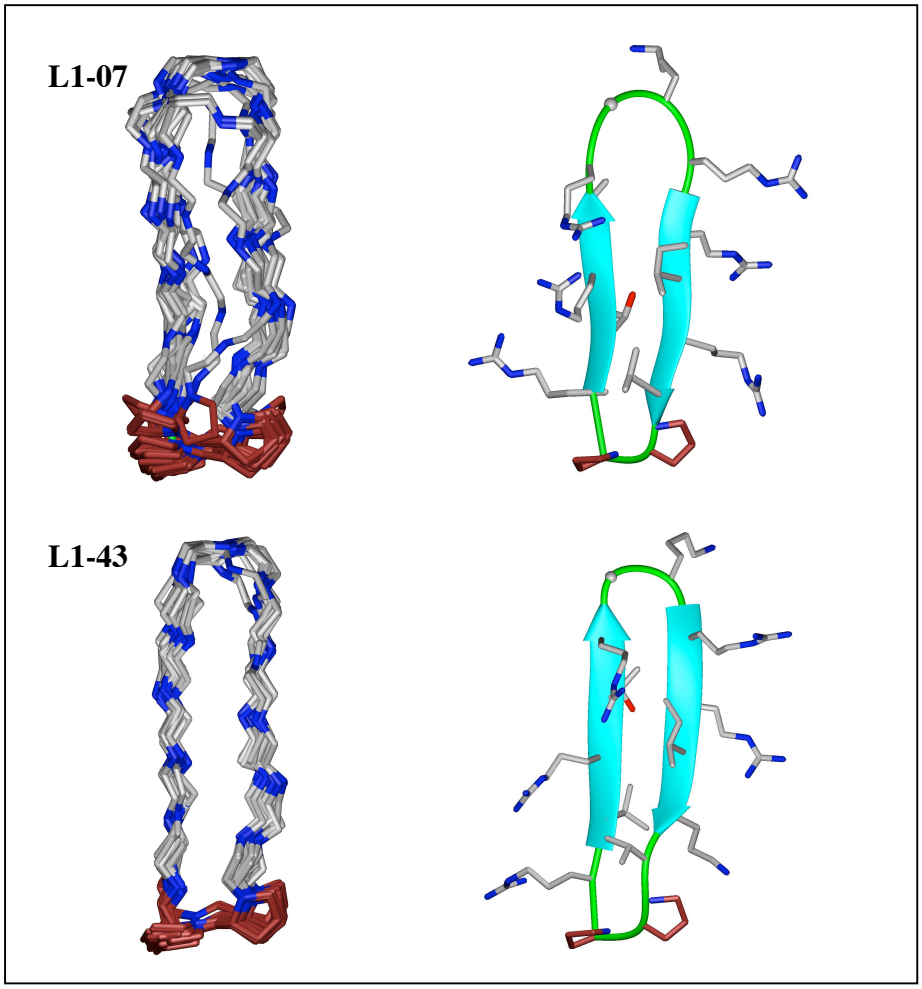



**Figure 27.** Left: Backbone superposition and representation of the final 20 NMR structures. Right: typical NMR structure with the ribbon representation of two strands for L1-07 and L1-43.

The temperature dependence of amide proton chemical shifts, as well as relative H/D exchange rates, provides complementary information about the involvement of amide NH groups in hydrogen bonding and/or their solvent accessibility.[104] Relatively small temperature coefficients (<5) were observed for residues 1, 3, 5, 8, 10 and 12 for all the studied peptides as expected for residues in cross-strand hydrogen bonding positions (Figure 28).

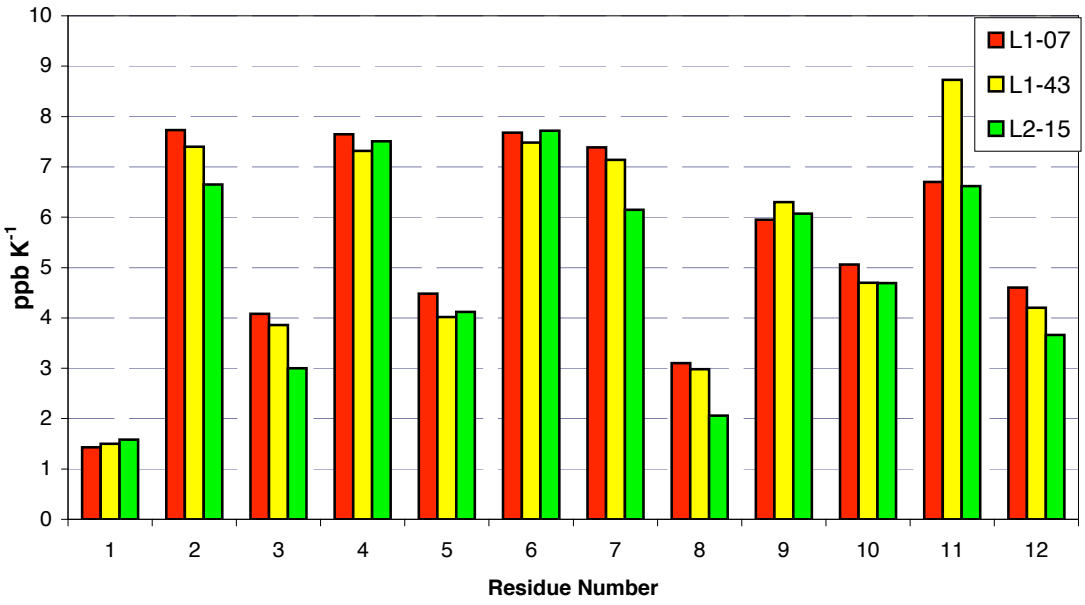



**Figure 28.** Comparison of NH temperature coefficients (in parts per billion per Kelvin) determined over range the 278-328K.

Further support for the strong network of cross-strand hydrogen bonding in L2-15 came from the H/D exchange experiment. Peptide bonds involved in H-bonding, or otherwise shielded from the solvent, can be identified by their relatively slow rate of H-D exchange. This effect is generally explained by the rate limiting breaking of the H-bond prior to exchange with the solvent.[105] Slowly exchanging amide protons in L2-15 were detected by monitoring the signal intensities of amide proton resonances after dissolution in $D_2O$ at 298K. The experiment showed that the exchange rates of amide NHs are significantly slowed for Arg1, Arg3 and in particular Ile10 and Ile12. The very slow exchange rates of Ile10 (half-time, $t_{1/2}$=32h) and Ile12 (half-time, $t_{1/2}$=13h) NHs might be due not only to their involvement in cross-strand H-bonding, but also due to shielding from the solvent by their aliphatic side chains (Figure 29 and Table 23).

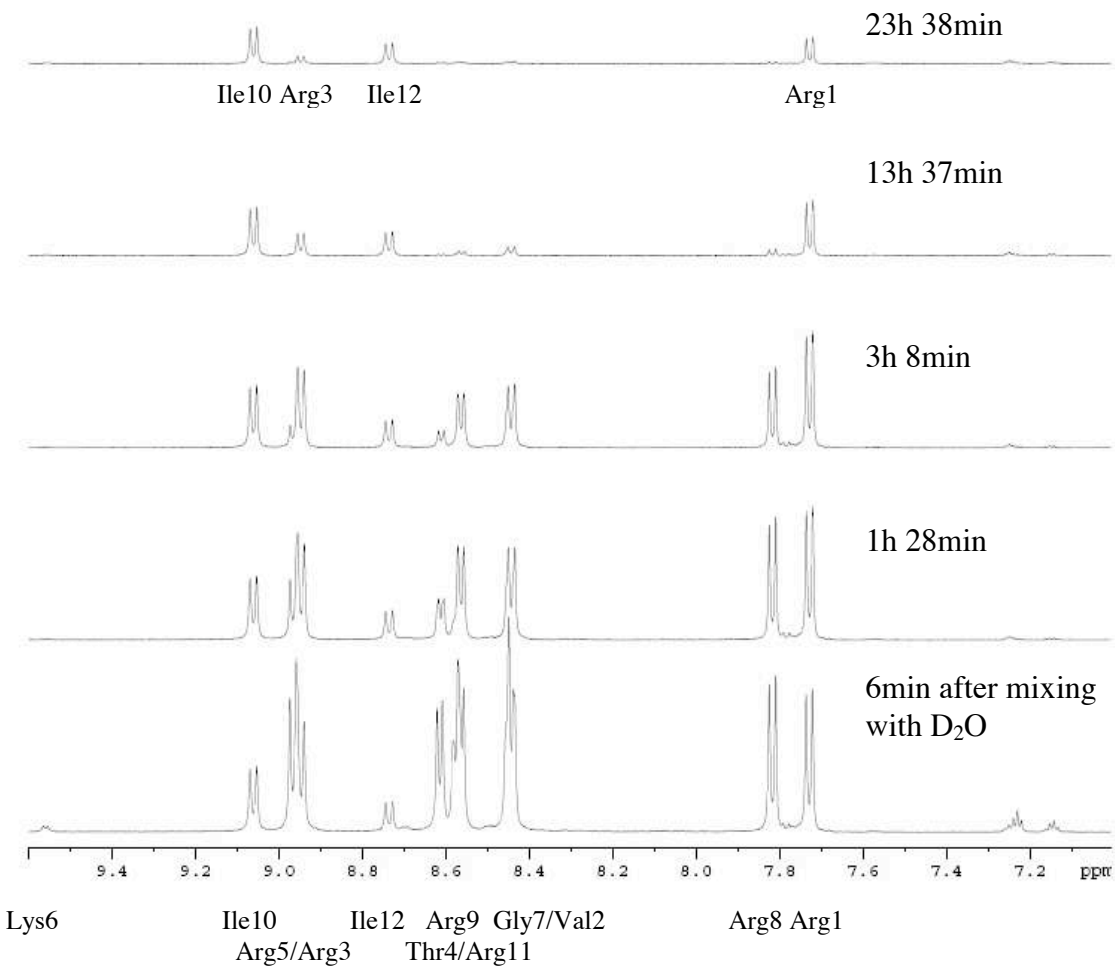



**Figure 29.** The H/D exchange experiment for L2-15 at 298K and pH=2.3. Hydrogen atoms of peptide bonds of Ile10, Arg3, Ile12 and Arg1 are the least susceptible to exchange in the experiment conditions.

| **L2-15** | Arg1 | Val2 | Arg3 | Thr4 | Arg5 | Lys6 | Gly7 | Arg8 | Arg9 | Ile10 | Arg11 | Ile12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **k [$min^{-1}$]** | 0,0013 | 0,0032 | 0,0021 | 0,0142 | 0,0097 | 0,2916 | 0,0035 | 0,0041 | 0,0116 | 0,0005 | 0,0043 | 0,0009 |
| **$\tau_{1/2}$ [min]** | 533,6 | 216,6 | 330 | 48,8 | 71,5 | 2,4 | 198,1 | 169,1 | 59,8 | 1926,6 | 161,2 | 770,3 |
| **H-bond** | S | M | M | W | W | W | M | M | M | S | M | S |

**Table 23**. H-D exchange rates and half-lives of peptide amide protons for L2-15. S-strong, M-medium and W-weak H-bonding, inferred from the exchange data.

The same set of experiments were performed for the selected "longer-loop" peptides exhibiting high affinity for HIV-1 TAR RNA, namely L4-01, L4-04, L4-07, L4-09 and L5-11 (Table 24).

**Table 24.** Sequences of the selected 18-amino acid peptidomimetics, whose conformations were studied in detail by NMR methods. Residues 1 and 16 are attached to the D-Pro-L-Pro template. The Cys residues in each peptide are linked by a disulfide bond.

| **Mimetic** | **Position** | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | ***13*** | **14** | ***15*** | **16** |
| **L4-01** | R | V | R | *C* | R | T | R | G | K | R | R | R | *C* | A | R | V |
| **L4-04** | R | V | R | *C* | R | K | R | $^{D}$R | G | Q | T | R | *C* | I | R | I |
| **L4-07** | R | V | R | *C* | R | R | K | G | P | G | Q | R | *C* | I | R | I |
| **L4-09** | R | *C* | R | V | R | Q | R | K | G | R | A | R | K | I | *C* | I |
| **L5-11** | R | V | R | *C* | R | Q | R | K | G | R | R | A | *C* | I | R | I |

The loop was extended to 16 amino acids stabilized by the template and by the disulfide bridge at the positions 2 and 15 for L4-09 and 4 and 13 for the other peptides. It was crucial to provide the experimental data to prove the conformational stability of the extended loop. The NMR experiments for 18-mer peptides were performed identically to the measurements for shorter peptides described above. The statistics important for the structural calculations are listed in Table 25 (see Appendix 6 for $^{1}$H-NMR assignments).

**Table 25.** Experimental distance restrains and statistics for the final 20 NMR structures calculated for L4-01, L4-04, L4-07, L4-09 and L5-11.

| **Mimetic** | **L4-01** | **L4-04** | **L4-07** | **L4-09** | **L5-11** |
|---|---|---|---|---|---|
| **NOE upper-distance limits** | 149 | 167 | 120 | 183 | 152 |
| Intraresidue | 59 | 59 | 39 | 62 | 62 |
| Sequential | 61 | 71 | 60 | 78 | 67 |
| Medium and Long-Range | 29 | 37 | 21 | 43 | 23 |
| **Residual Target Function Value (Å$^2$)** | 0.67±0.00768 | 0.67±0.12 | 0.74±0.0088 | 0.8±0.0095 | 0.71±0.097 |
| **Mean rmsd values (Å)** | | | | | |
| All backbone atoms | 1.09±0.47 | 1.53±0.60 | 1.63±0.53 | 1.48±0.98 | 1.40±0.41 |
| All heavy atoms | 2.77±0.5 | 3.24±0.81 | 3.15±0.65 | 3.26±1.54 | 3.16±0.56 |
| **Residual NOE violations** | | | | | |
| Violation > 0.2 Å | none | 5 | 11 | 7 | 7 |
| Maximum violation (Å) | none | 0.30 | 0.34 | 0.34 | 0.31 |

Structures of the peptides studied are shown in Figure 30 and Figure 31. Structures of L4-04, L4-07, L4-09 and L5-11 exhibit less convergence than the structure of L4-01 (seen also in increased backbone rmsd values, Table 25). This is due to the fact that visualised structures based on the calculations contain several NOE violations.

However, these violations are a minor component of all the NOE restraints and they are randomly distributed in all the 20-lowest energy final structures.

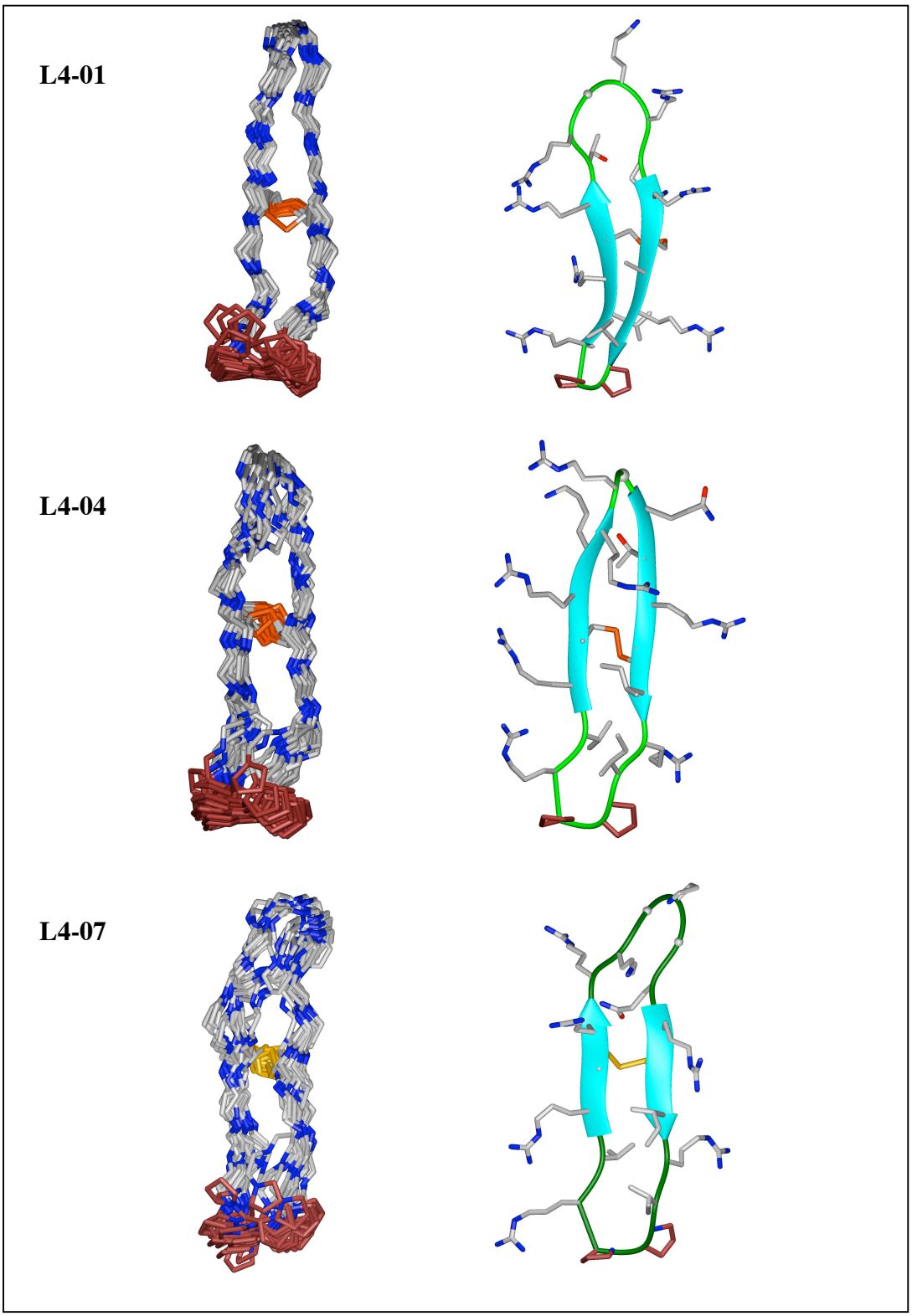



**Figure 30.** Left: Backbone superposition and representation of the final 20 NMR structures. Right: typical NMR structure with the ribbon representation of two strands for L4-01, L4-04 and L4-07.

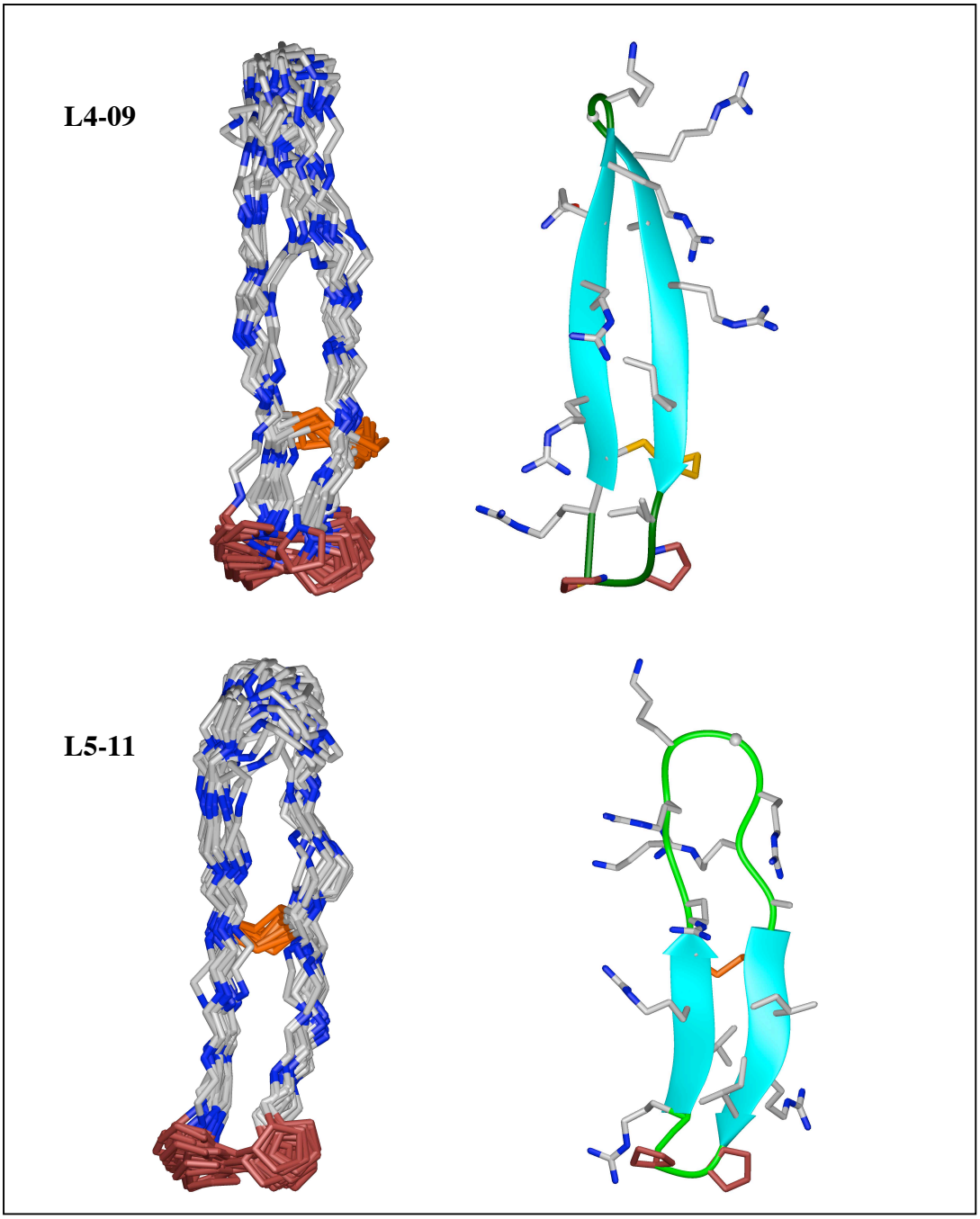



**Figure 31.** Left: Backbone superposition and representation of the final 20 NMR structures. Right: typical NMR structure with the ribbon representation of two strands for L4-09 and L5-11.

A comparison of the chemical shifts observed in these cyclic peptides with the random coil values ($\Delta\delta H\alpha$ >0.1 ppm) indicates an extended β-sheet conformation for residues 1-7 and 10-16 in all the studied peptides (Figure 32).

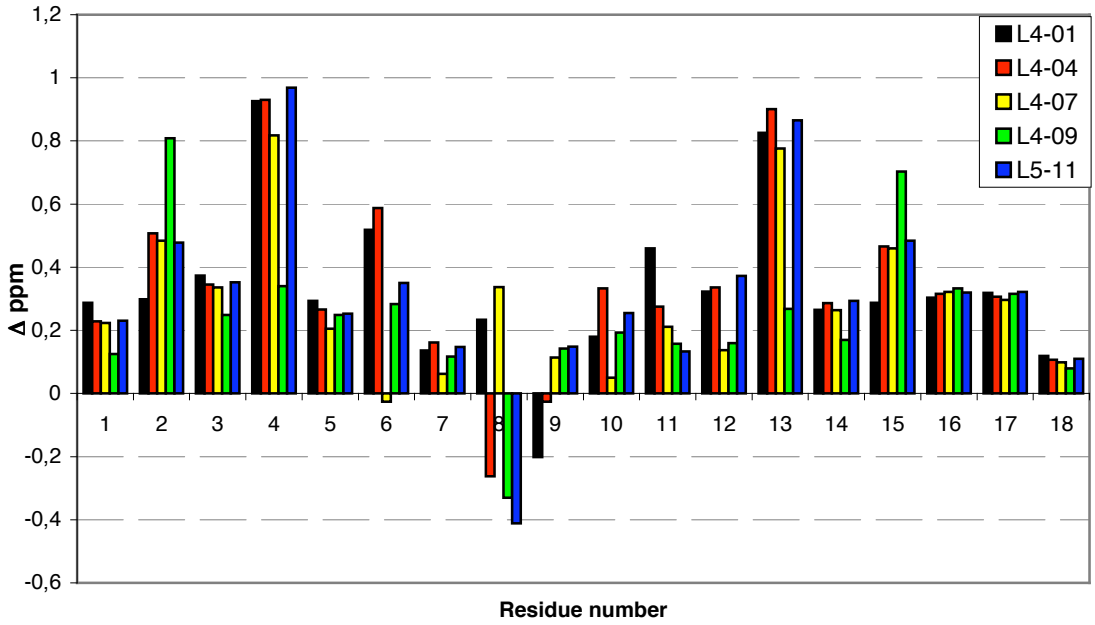



**Figure 32.** Hα chemical shifts relative to random coil chemical shifts of residues in peptidomimetics L4-01 (black), L4-04 (red), L4-07 (yellow), L4-09 (green) and L5-11 (blue). Residues 17 and 18 refer to D-Pro-L-Pro template.

The absolute values of temperature coefficients are shown in Figure 33. As expected, amide protons at positions 1, 3, 5 and 10 and to a smaller extent at positions 7 and 16 exhibit smaller changes in chemical shifts upon temperature increase, which indicates the involvement of these protons in stable hydrogen bonds.

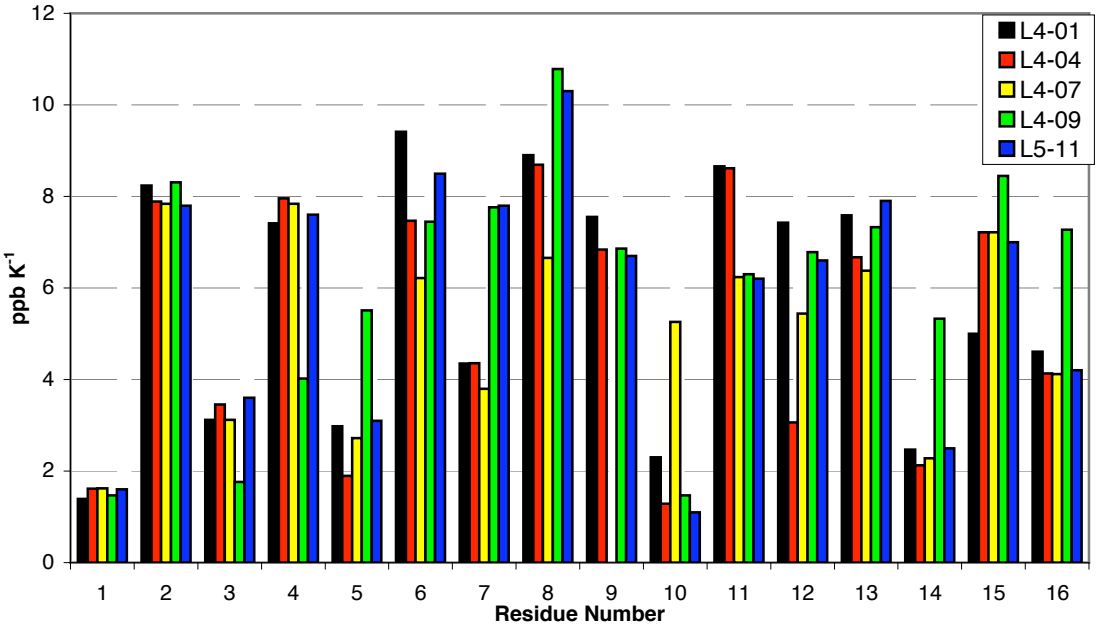



**Figure 33.** Comparison of NH temperature coefficients (in parts per billion per Kelvin) determined over the range 278-328K.

The H-D exchange experiments confirm the above mentioned results and the stability of the H-bonding pattern, especially in the smaller loop enclosed by the template and the disulfide bridge. Positions 1, 3 and 14, 16 exhibit very slow exchange rates, with half-times for exchange of up to 7 days or longer for L4-04 (Table 26). These values suggest that this peptide has an extremely stable conformation stabilized by persistent hydrogen bonds. The extremely slow exchange rates of amide protons of Ile14 and Ile16 in L4-04 might be due to shielding of peptide bonds by long, branched and unpolar side chains in addition to strong H-bonding.

| **L4-01** | **Arg1** | **Val2** | **Arg3** | **Cys4** | **Arg5** | **Thr6** | **Arg7** | **Gly8** | **Lys9** | **Arg10** | **Arg11** | **Arg12** | **Cys13** | **Ala14** | **Arg15** | **Val16** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| k $[min^{-1}]$ | 0,0015 | 0,125 | 0,0075 | >0.4 | >0.4 | 0,11 | 0,05 | >0.4 | >0.4 | 0,17 | 0,33 | 0,01 | >0.4 | 0.00039 | 0.2 | 0.00059 |
| $\tau_{1/2}$ [min] | 672 | 8 | 133 | >2.5 | >2.5 | 9 | 20 | >2.5 | >2.5 | 6 | 3 | 99 | >2.5 | 2552 | 5 | 1700 |
| H-bond | S | W | M | W | W | W | W | W | W | W | W | M | W | S | W | S |
| **L4-04** | **Arg1** | **Val2** | **Arg3** | **Cys4** | **Arg5** | **Lys6** | **Arg7** | **$^{D}$Arg8** | **Gly9** | **Gln10** | **Thr11** | **Arg12** | **Cys13** | **Ile14** | **Arg15** | **Ile16** |
| k $[min^{-1}]$ | 8.11e-05 | 0,0025 | 7.1e-05 | >0.2 | >0.4 | 0,067 | 0,05 | >0.5 | >0.25 | 0,0033 | 0,0625 | 0,0014 | >0.17 | 6.65e-05 | 0.0033 | 6.96e-05 |
| $\tau_{1/2}$ [min] | 8546 ~6days | 277 | 9763 ~7days | >5 | >3 | 10 | 20 | >2 | >4 | 210 | 11 | 488 | >6 | 10415 >7days | 207 | 9948 ~7days |
| H-bond | S | M | S | W | W | W | W | W | W | W | W | M | W | S | W | S |
| **L4-07** | **Arg1** | **Val2** | **Arg3** | **Cys4** | **Arg5** | **Arg6** | **Lys7** | **Gly8** | **Pro9** | **Gly10** | **Gln11** | **Arg12** | **Cys13** | **Ile14** | **Arg15** | **Ile16** |
| k $[min^{-1}]$ | 7.59e-04 | 0,0082 | 0.0034 | ~0.23 | 0.053 | 0,069 | 0,046 | ~0.23 | - | 0,116 | 0,0630 | 0,0433 | ~0.23 | 0.0014 | 0.0108 | 1.204e-04 |
| $\tau_{1/2}$ [min] | 913 | 85 | 206 | ~3 | 13 | 10 | 15 | ~3 | - | 6 | 11 | 16 | ~3 | 495 | 64 | 5756 |
| H-bond | S | M | M | W | W | W | W | W | - | W | W | W | W | M | W | S |

**Table 26.** H-D exchange rates and half-lives of peptide amide protons for longer-loop peptides. S-strong, M-medium and W-weak H-bonding, inferred from the exchange rate data.

At this point it was of interest to generate models of the RNA-bound forms of these cyclic peptides. The structures of the free HIV-1 TAR$^{17-45}$ RNA$^{106}$ and its Tat-derived peptide-bound form$^{72}$ have been determined at high resolution using NMR spectroscopy. The model complexes between potent representatives of the 14-mer peptides (L2-15) and 18-mer peptides (L4-04) were generated by superposition of the peptide structure determined in solution onto the previously solved structure of HIV-1 TAR RNA.

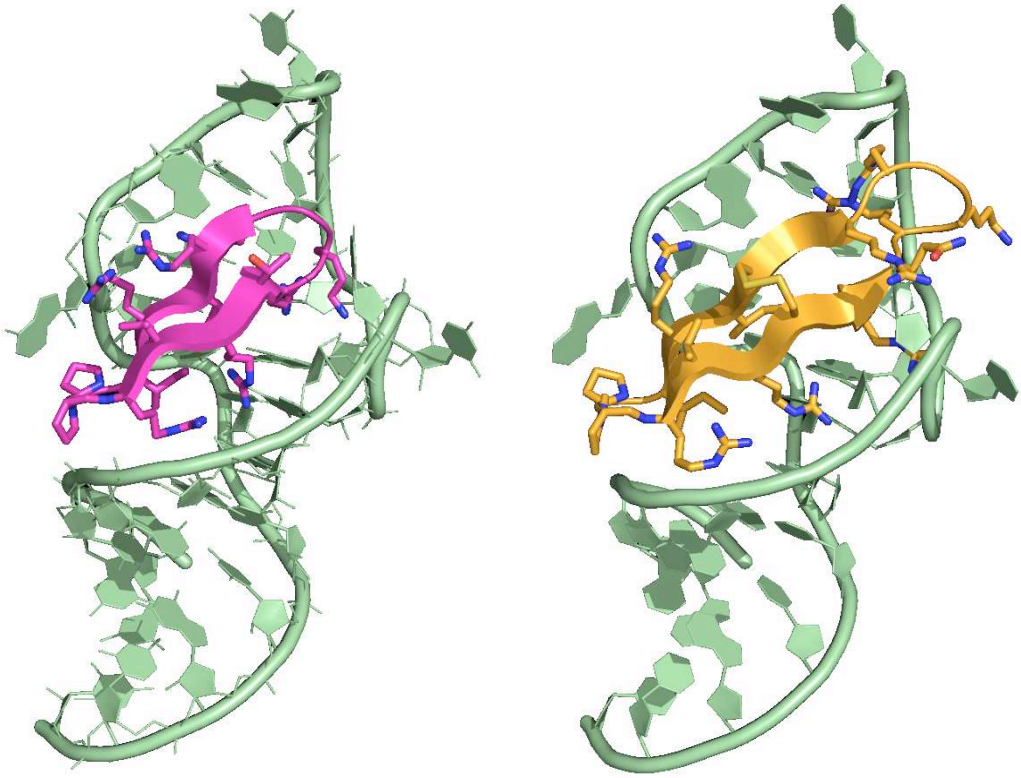


**Figure 34.** Model of complex: left: L2-15-HIV-1 TAR RNA; right: L4-04-HIV TAR RNA. The NMR-solved structures of L2-15 and L4-04 are superimposed on the HIV-1 TAR RNA NMR structure.

The superposition of both peptides together on the HIV-1 TAR RNA shows that by extending the peptide chain, it is likely that more contacts between the RNA and the peptide would be present (Figure 34). The tip of the loop of the peptide is in close proximity to the tip of the loop of the RNA, and this fact has a high importance when considering the formation of the ternary complex between cyclin T1, Tat protein and RNA, which is necessary for the successful reactivation of RNA polymerase II.

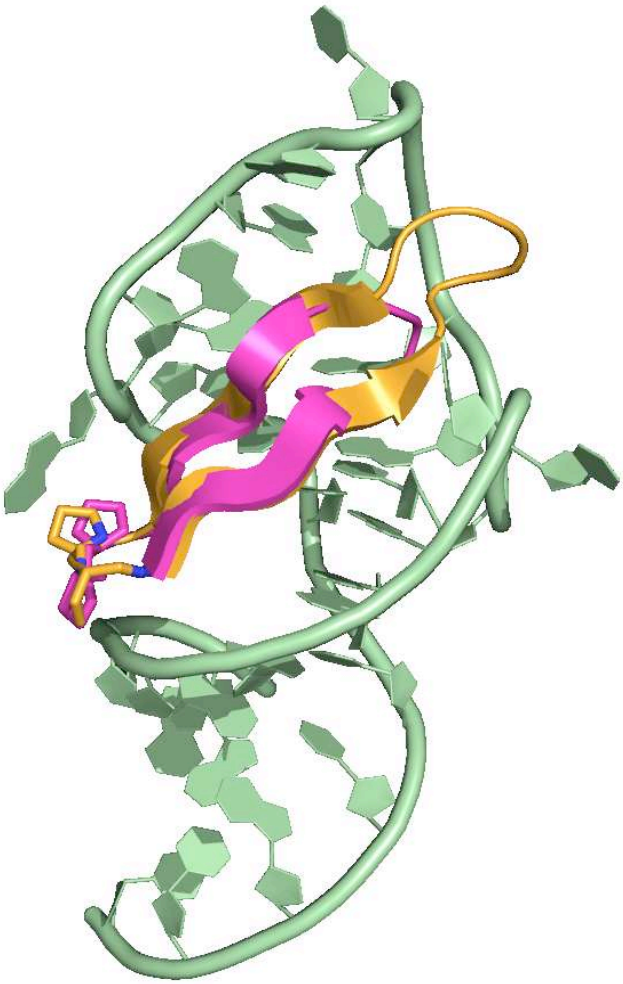


**Figure 35.** Comparative model of the complexes between 14-mer peptide, L2-15 (pink), and 18-mer peptide, L4-04 (orange) with HIV-1 TAR RNA.

The complexes between L2-15 and L4-04 and HIV-1 TAR RNA are currently under study by NMR spectroscopy in cooperation with Prof. G. Varani.

## 2.2.6. CD analysis

CD spectroscopy provides preliminary qualitative information on the secondary structure of the peptides under study. Peptides assuming a high proportion of β-sheet secondary structure typically show spectra with a minimum at 218nm (π–>π* transition) and a maximum at 196nm (n->π* transition).

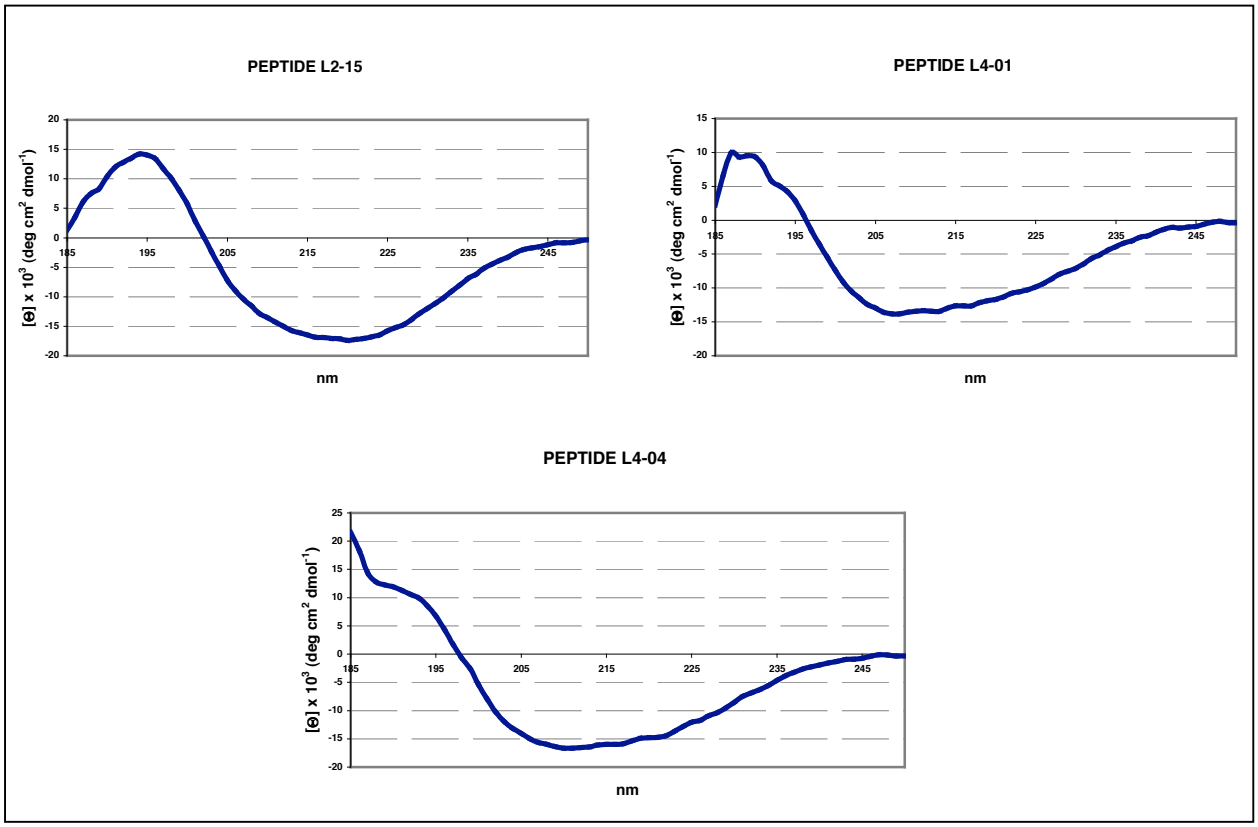



**Figure 36.** CD spectra of selected peptides: 14-mer L2-15 and 18-mer L4-01 and L4-04. Samples prepared in 10 mM phosphate buffer (pH=7.4), were measured in 0.1 mm quartz thermostatable cuvette at 25°C, using JASCO J-715 spectropolarimeter. A scan speed 5 nm/min, 2 seconds response and a 2 nm spectral band width were used. CD spectra are presented after subtraction of a blank containing only buffer and the optical activity is reported in units of mean residue ellipticity (deg $cm^2$ $dmol^{-1}$). Spectra presented as an average of five scans and results are given for a range 185-250 nm.

All three peptides under study show a shape of the CD spectrum characteristic of a β-sheet conformation. L2-15 shows a maximum intensity of 14200 deg $cm^2$ $dmol^{-1}$ at 194 nm and a minimum at 220 nm of -17400 deg $cm^2$ $dmol^{-1}$. These results are very close to the standard curves of molecules with 100% content of β-sheet secondary structure.[107] The longer loop peptides exhibit a shift in the wavelength of the minima and maxima indicating a slightly lowered content of a regular β-sheet secondary structure but they still share high similarities to the standard curve.

### 2.2.7. Cell permeability studies

Having found active Tat peptidomimetics, that bind tightly to the TAR RNA and therefore are potential antiviral agents, it is important to prove that the peptides can be delivered into cells, where their inhibitory action takes place. During the last decade, several peptides and proteins have been found to translocate through cellular membranes in a process called "protein transduction". These translocatory proteins share several common biophysical features that may be related to their apparent permeability. All appear to be taken up into the cytoplasm in a punctuated manner without accumulation in the nucleus, despite earlier experiments with fixed cells, which due to fixation artefacts appeared to show nuclear accumulation.[108] Each of them have a highly basic region that seems to mediate the ability of these proteins to bind to polyanions such as heparin/heparan sulfate, polysialic acid, and nucleic acids. These short sequences typically comprise less than 20 amino acids, which are highly rich in basic residues, and are called "protein transduction domains (PTDs)" or "cell penetrating peptides" (CPPs). These PTDs are responsible for the ability of the proteins to translocate across the plasma membrane. HIV Tat was one of the first membrane-permeable arginine-rich proteins to be discovered in the late 1980's.[109] In 1997, Vivès et al. reported that the arginine-rich segment in the Tat protein (positions 48–60) is the critical component for translocation.[110] Soon other arginine rich peptides were discovered that can internalise efficiently through the cell membrane, including the RNA-binding segments derived from HIV-1 Rev, the flock house virus (FHV) coat, the brome mosaic virus (BMV) Gag, and the human T-cell lymphotrophic virus (HTLV)-II Rex protein.[111,112] There are no sequence similarities amongst them, except that these segments contained more than seven arginine residues. As the number of arginines in the segment becomes smaller, the internalization of the peptides becomes less efficient. Peptides containing less than five arginines lose the ability to transduce across cell membranes.[113-115]

In the recent years, the translocation mechanism of major CPPs has been controversially discussed and is not yet fully understood. At early stages of such studies, it had been assumed that peptides such as Tat are delivered into the cell via direct, energy independent mechanisms and accumulate mainly in the nucleus.[110] Further studies proved this assumption to be wrong and the misleading results were explained by cell fixation artefacts and other experimental shortcomings. Recent studies

with life cells have demonstrated an energy and concentration dependent endocytic uptake mechanism.[108,116,117]

BIV has been used in our studies as a model of the TAR RNA/Tat protein interaction. BIV Tat peptide has the same function as HIV Tat and similar properties, such as a high content of basic residues in the region responsible for RNA-binding. BIV Tat permeability has not yet been studied, but it could be presumed that it also can translocate through membranes. Hence it was decided to investigate the cell permeability of the BIV and HIV Tat mimetics described in the previous section.

Peptidomimetics that target intracellular interactions must be efficiently delivered into the cell to perform their inhibitory action. BIV2 and L1-07 were chosen as representative examples of the β-hairpin peptidomimetics of BIV Tat and HIV-1, and the corresponding Tat-derived peptides were used as positive controls. The chosen peptides were fluorescently labelled in order to observe their uptake and cell localisation in HeLa cells by means of confocal microscopy (Table 27).

A selection of three fluorescent probes was used: **i**) 5(6)-carboxyfluorescein ($\lambda_a$= 494 nm and $\lambda_e$= 521 nm), **ii**) 5(6)-carboxyfluorescein diacetate ($\lambda_a$= 494 nm and $\lambda_e$= 521 nm), **iii**) 5(6)-carboxytetraethylrhodamine ($\lambda_a$= 554 nm and $\lambda_e$= 580 nm) (Figure 37).

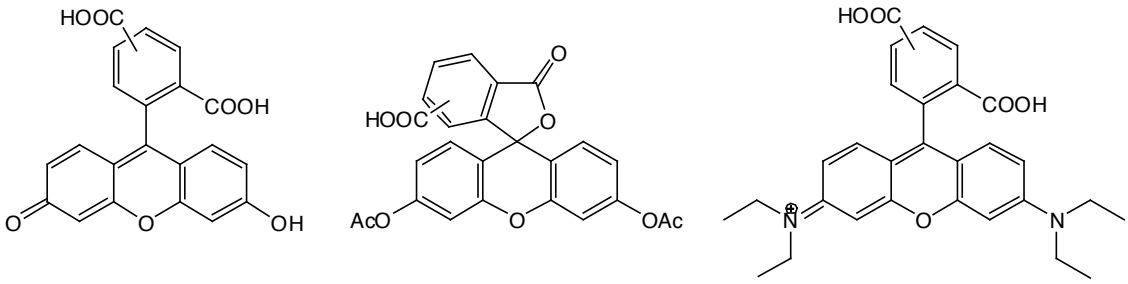



**Figure 37.** Fluorescent probes used to label the selected peptides. Left: 5(6)-carboxyfluorescein, middle: 5(6)-carboxyfluorescein diacetate, right: 5(6)-carboxytetraethylrhodamine.

5(6)-Carboxyfluorescein is a very commonly used probe for permeability studies, however, it is negatively charged in physiological conditions. To evaluate if the charge of the probe has any influence on cell permeability we used in addition 5(6)-carboxyfluorescein diacetate, which is non-polar and only after entering the cell is converted into the highly fluorescent dianionic 5(6)-carboxyfluorescein[118] by intracellular esterase(s) (Figure 38).

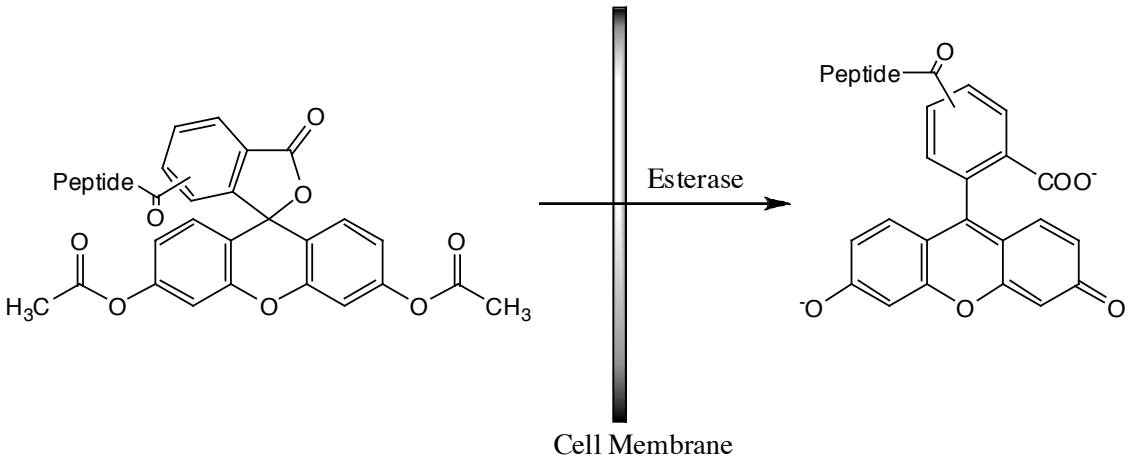



**Figure 38.** Mechanism of the conversion of non-polar 5(6)-carboxyfluorescein diacetate into ionic 5(6)-carboxyfluorescein.

### 2.2.7.1. Synthesis of labelled peptides for cell permeability studies

- **Synthesis of the linear labelled peptides**

The positive controls for the permeability studies were chosen according to literature precedence: **i**) HIV-1 Tat arginine rich domain between Gly38 and Gln60 has been shown to be critical for permeability,[110] **ii**) BIV Tat arginine rich domain between Ser65 and Lys82, which is assumed here to have the same permeability characteristic as the corresponding HIV Tat domain (Table 27).

The linear Tat peptides were selectively labelled on the resin with 5(6)-carboxyfluorescein or 5(6)-carboxyrhodamine at either the free N-terminal amino group, before other protecting groups were removed, or at the side chain of the Lys82 in BIV Tat after selective removal of an Alloc-protecting group, and prior to removal of other side chain protecting groups. The free carboxyl groups in the fluorophores were coupled to the amino group of the peptide using standard coupling methodology with HATU/HOAt. The fluorescent labels performed very well in the solid phase peptide synthesis and they were stable in the standard final deprotection conditions. The labelled linear peptides were isolated in good yield after HPLC purification, and were characterized by MS (see experimental part 5.5.4 for synthesis and analytical data).

- **Synthesis of the labelled cyclic peptides**

Labelling of the cyclic peptides requires an additional functionalisation of a chosen amino acid, since simply attaching it to the amine group of any lysine side chain present in the peptide may affect the binding affinity, as well as the cell permeability. Based on the structure of the BIV2/TAR RNA complex (Figure 39) D-Pro was chosen as a site for functionalisation and labelling, as it makes no direct contact with the RNA and therefore labelling at this residue should have minimal effects on binding. A hydrazino group at the 4-position of this D-Pro would provide a strongly nucleophilic site at which a selective ligation to the fluorophore could be performed. The first step, therefore, was the synthesis of an appropriately protected (2*R*,4*S*)-4-hydrazino-D-proline.

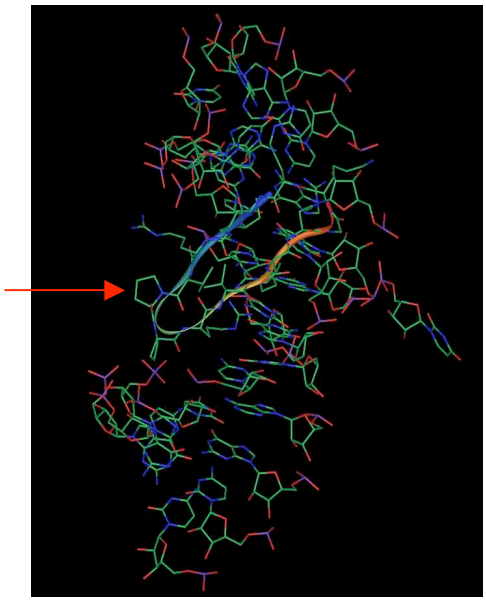


**Figure 39**. NMR-structure of the complex of BIV2-BIV TAR RNA. The D-Pro residue sticking out into solution is indicated with an arrow.

| Peptide | Sequence | Origins of the peptide |
|---|---|---|
| Fl-BIV-Tat | Fl-SGPRPRGTRGKGRRIRRK-$NH_2$ | BIV Tat protein residues 65-82 |
| Rh-BIV-Tat | H-SGPRPRGTRGKGRRIRRK[Rh]-$NH_2$ | BIV Tat protein residues 65-82 |
| Fl-HIV-Tat | Fl-GRKKRRQRRRPPQ-$NH_2$ | HIV Tat protein residues 48-60 |
| Rh-BIV2$^{h}$ | cyclo(RVRTRGKRRIRV$^{D}$P[Rh]P) | earlier work |
| Fl-BIV2$^{h}$ | cyclo(RVRTRGKRRIRV$^{D}$P[Fl]P) | earlier work |
| Rh-L1-07$^{h}$ | cyclo(RTRTRGKRRIRV$^{D}$P[Rh]P) | L1 library |
| Fl-L1-07$^{h}$ | cyclo(RTRTRGKRRIRV$^{D}$P[Fl]P) | L1 library |
| Fl-Ac-BIV2$^{h}$ | cyclo(RVRTRGKRRIRV$^{D}$P[Fl-Ac]P) | earlier work |
| Fl-Ac-L1-07$^{h}$ | cyclo(RTRTRGKRRIRV$^{D}$P[Fl-Ac]P) | L1 library |

**Table 27.** Selection of peptides labelled with fluorescent probes, including linear positive controls derived from BIV and HIV-1 Tat arginine-rich domains and cyclic peptides BIV2 and L1-07. Fl=Fluorescein, Rh=Rhodamine, Fl-Ac=Fluorescein diacetate, $^{h}$=indication that the peptide contains the 4-hydrazino-D-Pro in the template

- **Synthesis of (2*R*,4*S*)-4-[*N*,*N*,*N'*-tris-(*tert*-butyloxycarbonyl)]hydrazino-1-[(*9H*-fluoren-9-yl)-methoxycarbonyl]-proline**

D-Proline was modified at position $C_\gamma$ by a fully Boc-protected hydrazine group. This should allow a site-specific derivatization of a fully deprotected peptide, and is compatible with solid-phase peptide chemistry.[119] The hydrazino group has been used previously, for example, in the chemoselective acylation of peptides by fatty acid succinimidyl esters, providing a valuable access to complex lipopeptides.[120]

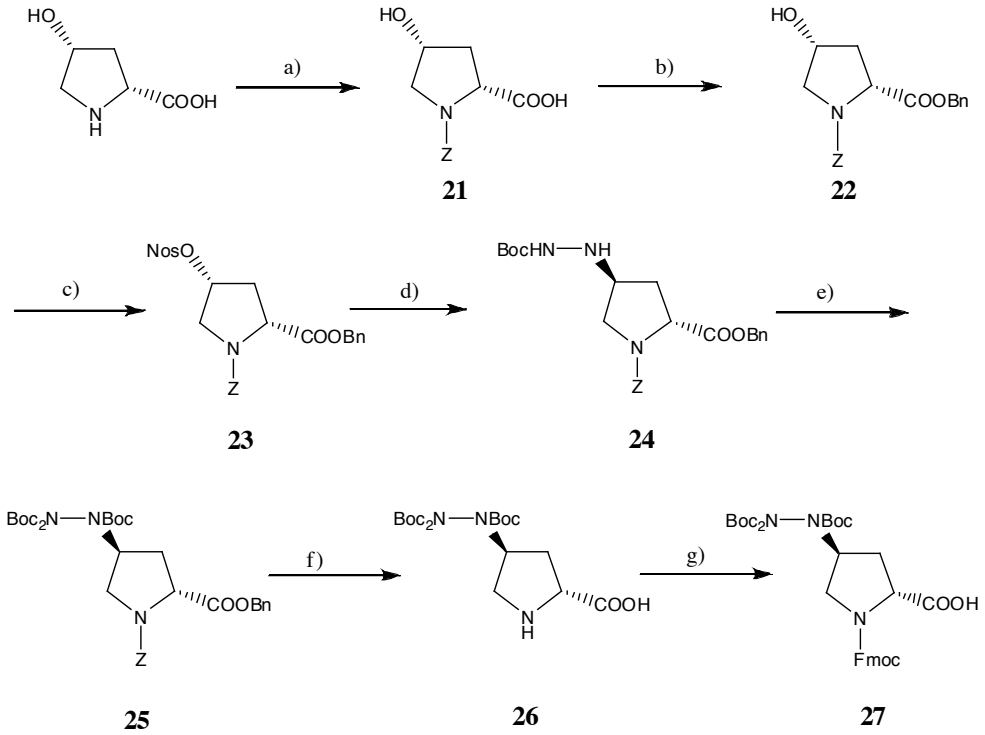



**Figure 40.** Scheme of synthesis of Fmoc-4-(*N*,*N*,*N'*-tris-Boc)-hydrazino-D-proline. Reagents a) $PhCH_2OCOCl$, NaOH, THF; 76%, b) BnBr, $K_2CO_3$, NaI, DMF; 97%, c) NosCl, TEA, DCM; 95%, d) $BocNHNH_2$, Dioxan; 79%, e) $Boc_2O$, DMAP, $NEt_3$, DCM; 75%, f) $H_2$, Pd/C 10%, MeOH; 75%, g) Fmoc-OSu, DIEA, DCM; 72%.

The synthesis of the Fmoc/Boc protected 4-hydrazino-D-proline (**27**) is shown in Figure 40, starting from (2*R*,4*R*)-4-hydroxy-D-proline. Protection of the secondary amino group (**21**) and conversion into the benzyl ester (**22**) were performed according to established methods.[121] Introduction of the nosyl group as an excellent leaving group (**23**) and its subsequent nucleophilic displacement by t-butylcarbazate (**24**) with a complete inversion of configuration was performed by slightly modifying the procedures of Hoffman et al.[122] and Lerner et al.[123] In the next step, the hydrazine

moiety was fully protected using di-*tert*-butyl dicarbonate to yield *N*-benzyl-4-[*N*,*N*,*N'*-tris-*tert*-butyloxycarbonylhydrazino]-proline benzyl ester (**25**).[119]

Hydrogenolysis of the fully protected derivative (**25**) removed the Nα-Z and Bn protecting groups and gave the desired product, 4-[*N*,*N*,*N'*-tris-*tert*-butyloxycarbonyl-hydrazino]-proline (**26**). The free amino group was then protected with Fmoc using Fmoc-OSu to deliver the amino acid (**27**) in overall yield of 22% in 7 steps, ready-to-use in solid phase peptide synthesis.

- **Preparation of the fluorescently labelled cyclic peptides**

For the synthesis of fluorescently labelled BIV2 and L1-07, the Fmoc/Boc protected hydrazino-D-proline (**27**) was coupled manually to L-Pro preloaded chlorotrityl chloride resin. To ensure the coupling was complete, the chloranil test was performed. Subsequently, the peptides were assembled on the peptide synthesizer using standard Fmoc chemistry. Cyclization and the final deprotection yielded the desired peptides (Figure 41), which were purified by HPLC and characterized by MS (see paragraph 5.5.4 for synthesis and analytical data).

The coupling of the fluorophores was performed in 50 mM citrate buffer pH=5.0-5.2. The $pK_a$ of the hydrazino group (~6)[120] should be about 4 units below that of the ε-amino group of lysine (~10), therefore performing the reaction under pH controlled conditions should preferentially lead to a single acylation at the hydrazine group. To improve coupling efficiency the fluorescent probes were activated as N-hydroxy succinimide esters, prepared freshly before reaction (reaction of fluorophore with N-hydroxysuccinimide and DIC) or purchased, if commercially available. To improve the solubility of the activated fluorophore, the probe was dissolved in DMF and then added to the buffered peptide solution.

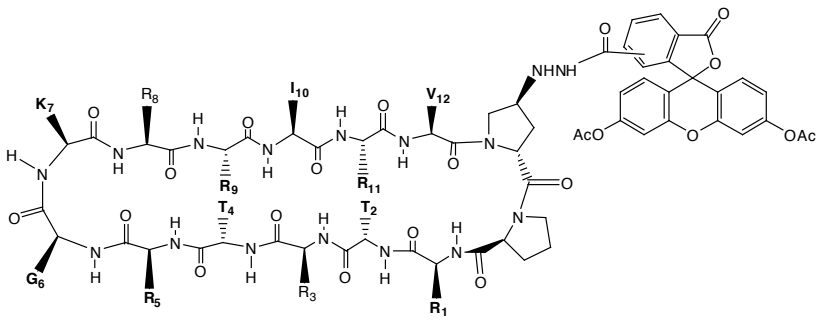



**Figure 41.** An example of the cyclic 14-amino acid peptide, L1-07$^h$, labelled with 5(6)-carboxyfluorescein diacetate at the hydrazine moiety in D-proline.

By carefully adjusting the content of DMF to solubilize the succinimidyl ester of the fluorophore, the time of the reaction, and the ratio of peptide to fluorescent probe, the yield of the labelling reaction could be improved from 6% to above 50%. Figure 42 shows typical HPLC chromatograms of crude reaction products, before and after optimization.

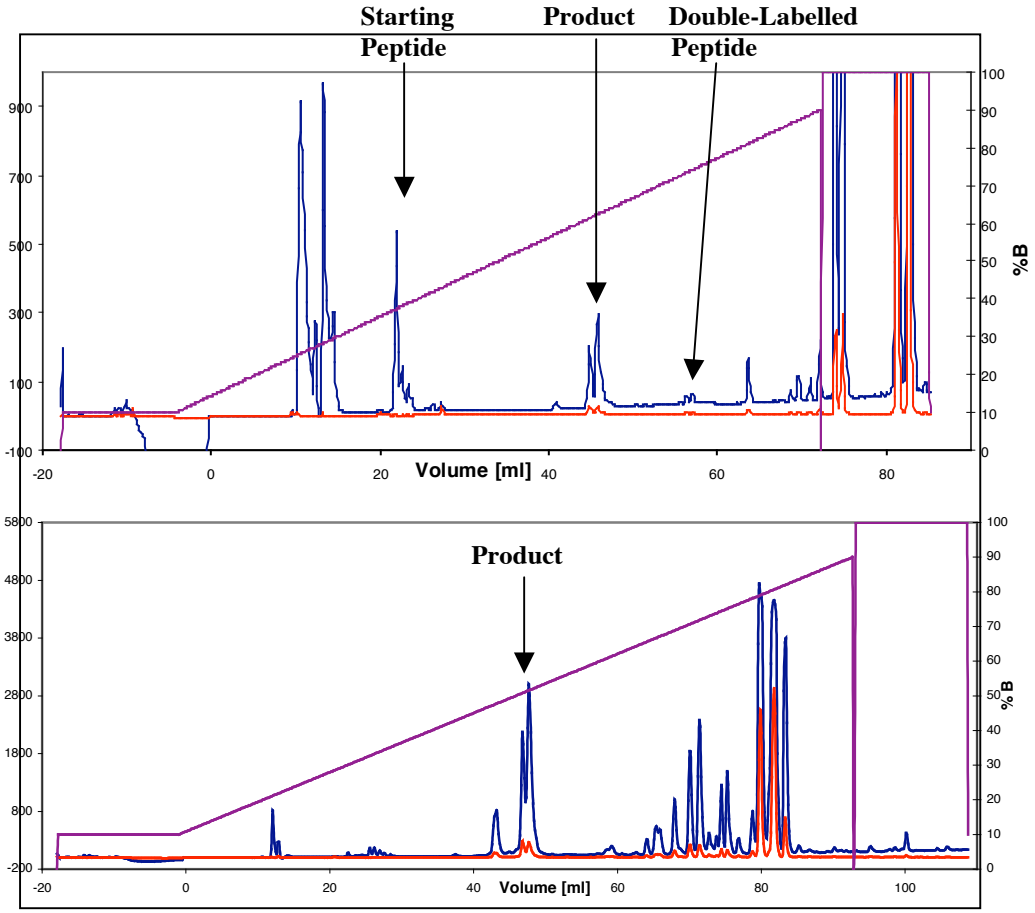



**Figure 42**. HPLC chromatograms of the reaction mixtures of labelling reaction with L1-07 and 5(6)-carboxyfluorescein-diacetate-succinimidyl ester. Semi-preparative Zorbax Eclipse column, $C_{18}$, gradient 10-90% MeCN in $H_2O$+0.1%TFA. Flow 5 ml/min. The optimised reaction (bottom) was performed in 50 mM citrate buffer (pH=5.0): DMF (5:1) over 5 days with 1.0 eq of the fluorophore.

#### 2.2.7.2. Cellular uptake of fluorescently labelled peptides into HeLa Cells

The distribution of the fluorescently labelled peptides was followed in living unfixed HeLa cells. The cells were incubated with the peptides at 37°C for 30 min and with Hoechst 33341 for 5 min for nuclear staining, and after washing with PBS buffer inspected immediately under the confocal microscope. A separate set of experiments was prepared using fixed cells (mild fixation with 3% paraformaldehyde, 2% sucrose in BPS buffer) to evaluate the influence of fixation on the peptide distribution. In the case of experiments with fixed cells, the cells after fixation were incubated for 5 min with DAPI for nuclear staining, washed with PBS and mounted on microscope slides. Hoechst 33341 and DAPI are known to form fluorescent complexes with natural double-stranded DNA and they were used to visualize the nucleus. The transmission images were recorded to visualize the shape of the cells and prove their health.

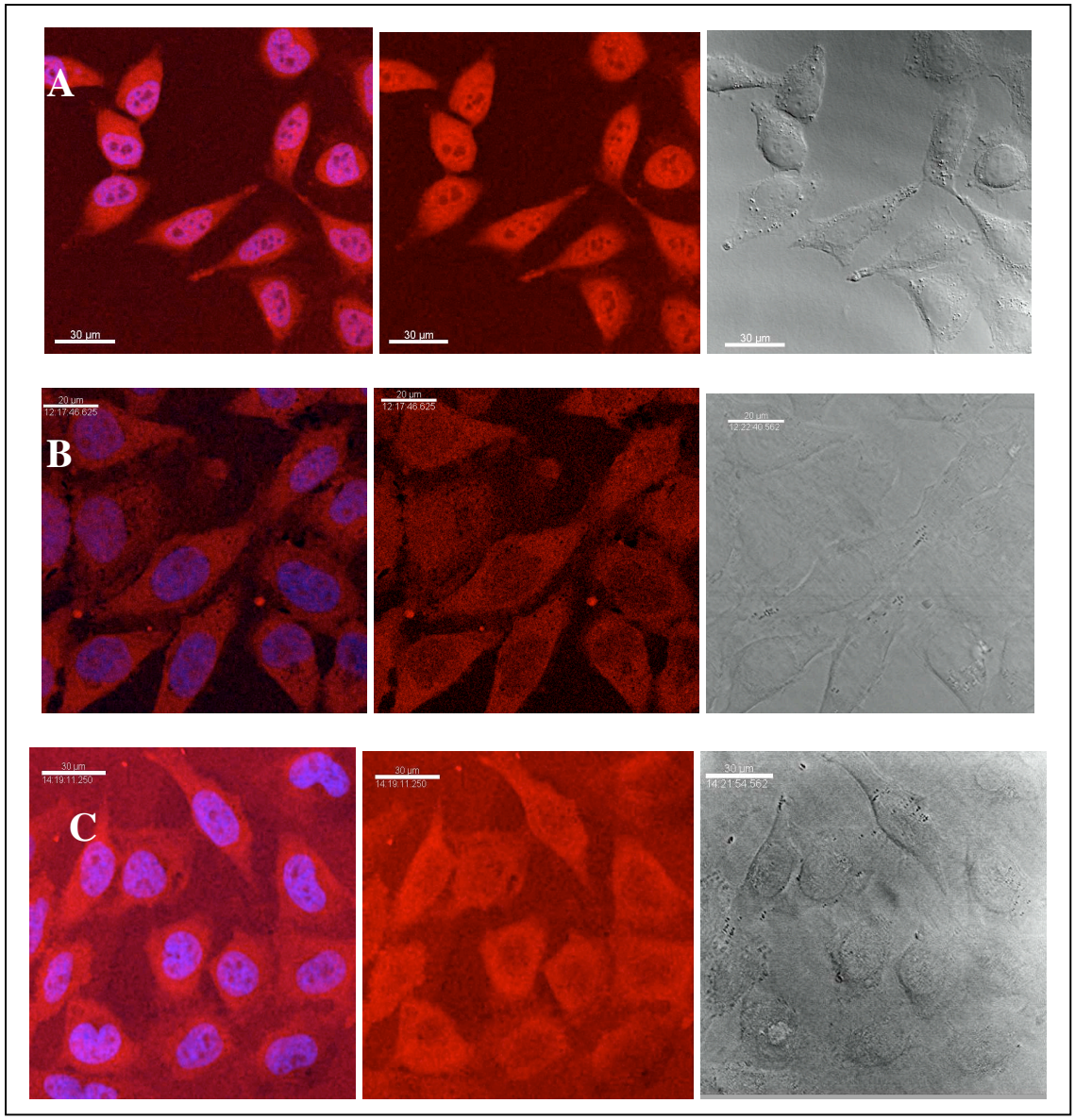



**Figure 43.** Cellular uptake of peptides labelled with **5(6)-carboxyrhodamine** in the **fixed** HeLa cells. $5\times10^4$ cells seeded on coverglass were incubated with 500 nM peptide solution for 30 min at 37°C. Fixation was with 3% paraformaldehyde, 2% sucrose in BPS. **Left:** Rhodamine-peptide (red) with DAPI (blue) for nuclear staining; **middle:** Rhodamine-peptide (red), **right:** transmission image of the cells. **A:** corresponds to Rh-BIV2$^h$, **B:** corresponds to Rh-L1-07$^h$, **C:** corresponds to Rh-BIV Tat.

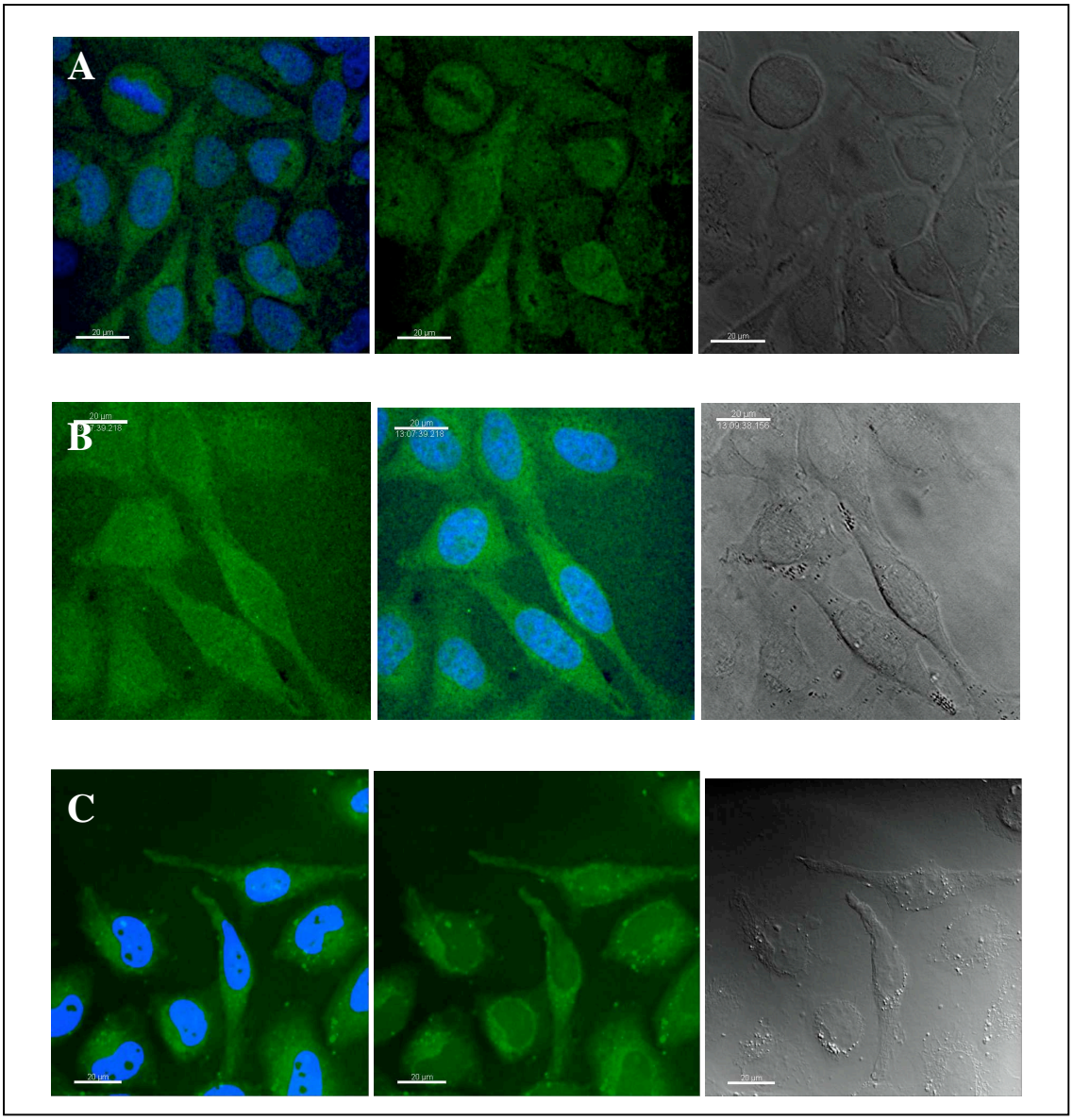



**Figure 44.** Cellular uptake of peptides labelled with **5(6)-carboxyfluorescein** in the **fixed** HeLa cells. $5\times10^4$ cells seeded on coverglass were incubated with 500 nM peptide solution for 30 min at 37°C. Fixation was with 3% paraformaldehyd, 2% sucrose in BPS. **Left:** Fluorescein-peptide (green) with DAPI (blue) for nuclear staining; **middle:** Fluorescein-peptide (green), **right:** transmission image of the cells. **A:** corresponds to Fl-BIV2$^h$, **B:** corresponds to Fl-L1-07$^h$, **C:** corresponds to Fl-HIV-Tat.

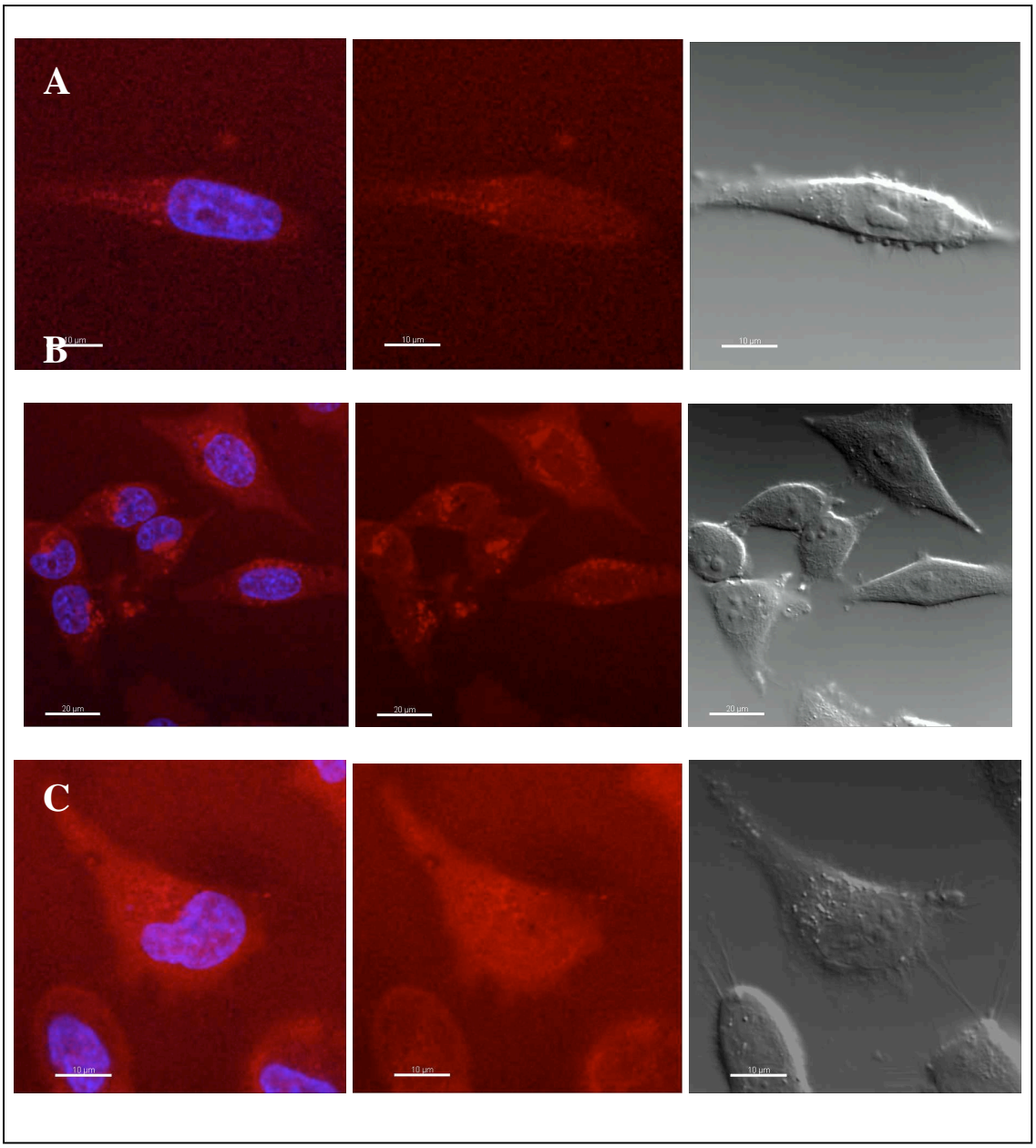



**Figure 45.** Cellular uptake of peptides labelled with **5(6)-carboxyrhodamine** in **live** HeLa cells. $8\times10^4$ cells seeded on coverglass were incubated with 1 µM peptide solution for 30 min at 37°C.
**Left:** Rhodamine-peptide (red) with Hoechst 33341 (blue) for nuclear staining; **middle:** Rhodamine-peptide (red), **right:** transmition image of the cells.
**A:** corresponds to Rh-BIV-Tat, **B:** corresponds to Rh-L1-07$^h$, **C:** corresponds to Rh-BIV2$^h$.

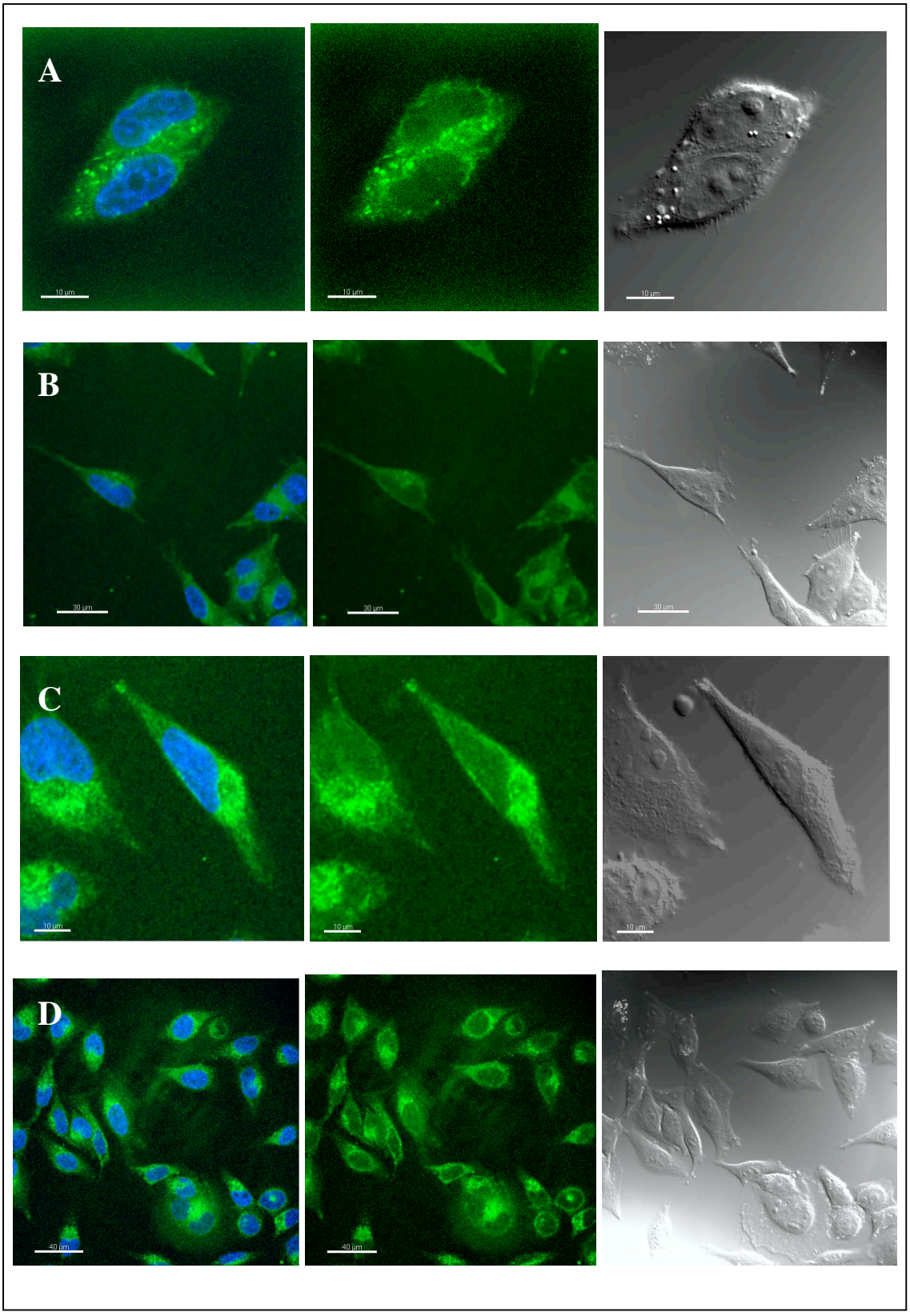



**Figure 46.** Cellular uptake of peptides labelled with **5(6)-carboxyfluorescein (diacetate)** in **live** HeLa cells. $8\times10^4$ cells seeded on coverglass were incubated with 1 μM peptide solution for 30 min at 37°C.
**Left:** Fluorescein (diacetate) -peptide (green) with Hoechst 33341 (blue) for nuclear staining; **middle:** Fluorescein (diacetate)-peptide (green), **right:** transmition image of the cells.
**A:** corresponds to Fl-HIV-Tat, **B:** corresponds to Fl-BIV-Tat, **C:** corresponds to Fl-Ac-L1-07$^h$, **D:** Fl-Ac-BIV2$^h$.

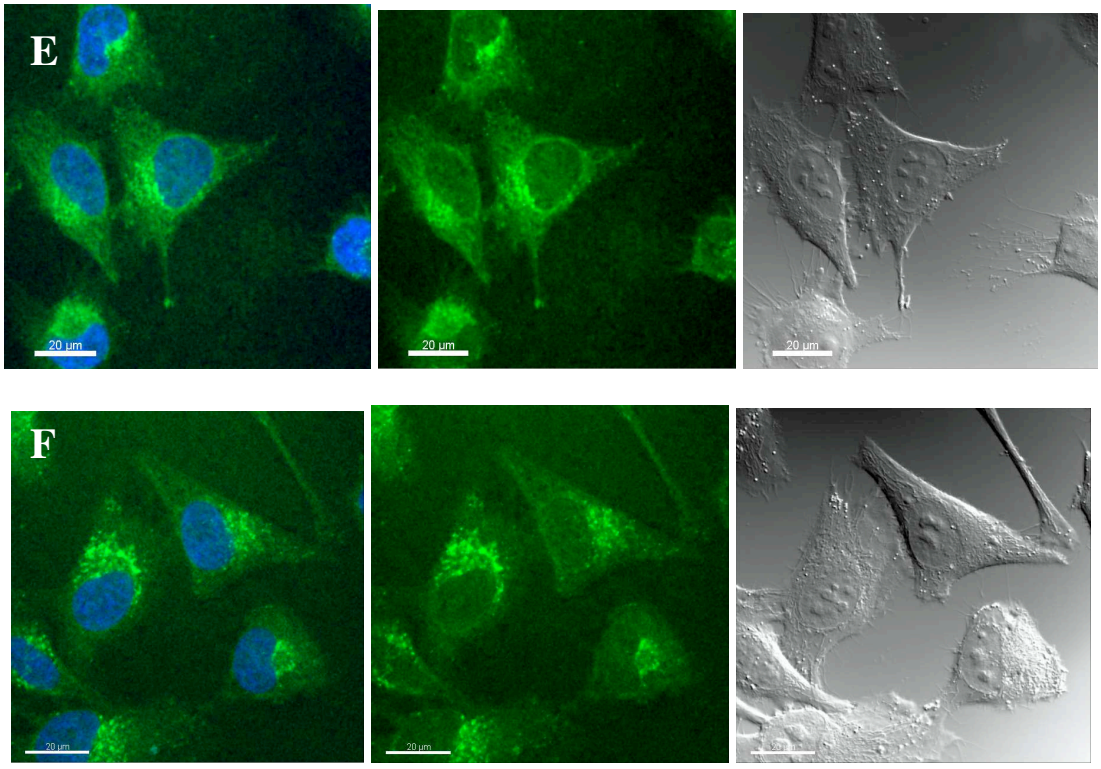



**Figure 47: (Continued)** **E:** corresponds to Fl-BIV2$^{h}$, **F:** corresponds to Fl-L1-07$^{h}$.

All peptides studied showed a punctuated vesicular-type, cytoplasmic distribution of the labelled peptides (Figure 43-Figure 47). The peptides were localized in vesicular structures that were predominantly on one side of the nucleus. BIV Tat and HIV Tat peptides were analysed as controls and showed the same fluorescence distribution as the cyclic peptides. No nuclear localization was observed, as frequently reported as a common artefact in experiments with fixed cells.[108,110] These observations are consistent with knowledge to-date on the endocytic translocation mechanism. The incubation time used in the experiments presented above was 30 min, although already after very short incubation times, such as 5 min, the internalisation was already taking place to visible levels. For the peptides labelled with fluorescein or fluorescein diacetate the quality of the images was mostly very good with a high resolution, while the images of the cells exposed to rhodamine most of the time could only be obtained with lower resolution. These problems might be due to spectroscopic differences between the two fluorophores.

### 2.2.8. Proteolytic and plasma stability of selected peptides

The stability of drug candidates in a biological environment is essential for maintaining acceptable drug concentrations and half-life in order to achieve desirable pharmacological effects. Unstable compounds tend to have rapid clearance and short half-lives, resulting in poor *in vivo* performance. Except for "pro-drugs", drug candidates undergoing rapid degradation by enzymes in plasma or tissue, generally have undesirable pharmacokinetic parameters and pose analytical challenges. Pharmaceutical companies tend to not advance the development of compounds that degrade rapidly, with the exception of prodrugs and special cases.

Cleavage by proteases is one of the main pathways for inactivation of peptides in biological settings. The major problem with therapy based on peptidic drugs is their susceptibility to extremely rapid enzymatic degradation after subcutaneous or oral administration. There is great demand for more stable peptidomimetic analogues, which would allow a patient-friendly way of administration, a reduction in dose and frequency of administration.

Selected Tat-derived peptides were studied here in two different assays to assess their rate of catabolism: **i**) in a trypsin digestion assay and, **ii**) in human and mouse plasma.

#### 2.2.8.1. Trypsin digestion assay

As proteases typically bind their substrates in a linear or an extended rather then folded conformation, inducing a stable β-hairpin secondary structure is expected to improve protease stability, leading to increased potency *in vivo*. To test this concept *in vitro*, selected peptides were treated with trypsin and the digestion process was analysed over time for up to 10 h. Trypsin is an endopeptidase (serine protease) that predominantly cleaves peptide chains at the carboxyl side of the amino acids lysine and arginine, except when either is followed by proline. The Tat peptidomimetics are rich in arginine and lysine residues. Therefore trypsin was chosen as a relevant proteolytic enzyme for the assay. **L2-15**, **L4-01** and **L4-04** were chosen for these studies, since their stable β-hairpin conformation has been confirmed by NMR studies (Paragraph 0). The linear version of the peptide L2-15 (**L2-15a**) was used to provide a comparative

study between linear and cyclic peptides in terms of their susceptibility to cleavage by trypsin. Trypsin activity was determined by methods published first by Erlanger[124] and modified by Temporini.[125]

| Peptide | Half-life $\tau_{1/2}$ [min] |
|---|---|
| L2-15a | 1.5±0.6 |
| L2-15 | 21±1 |
| L4-01 | 12.6±0.3 |
| L4-04 | 360±15 |

**Table 28.** The results of the proteolytic digestion of peptides by trypsin. Peptides (0.33 mg/ml) were incubated with trypsin (2 μg/ml) in 50 mM Tris buffer (pH=8.2) containing 10 mM $CaCl_2$. Reaction aliquots were taken at suitable time points. Digestion was stopped by addition of 30% AcOH. Samples were analysed by HPLC with L-Tryptophan (0.16 mM) as an internal control (method described by Wang et al.[126])

As expected, the linear control peptide, L2-15a, is highly susceptible to cleavage by trypsin and was completely degraded within a few minutes under the assay conditions. The cyclic compounds L2-15 and L4-01 were completely digested in under 60 min, while L4-04 was unexpectedly resistant to trypsin digestion and the complete degradation was detected only after more than 2 days. The half-lives ($\tau_{1/2}$) of the peptides under the assay conditions are given in Table 28. As an example, the data points for L4-01 are shown in Figure 48.

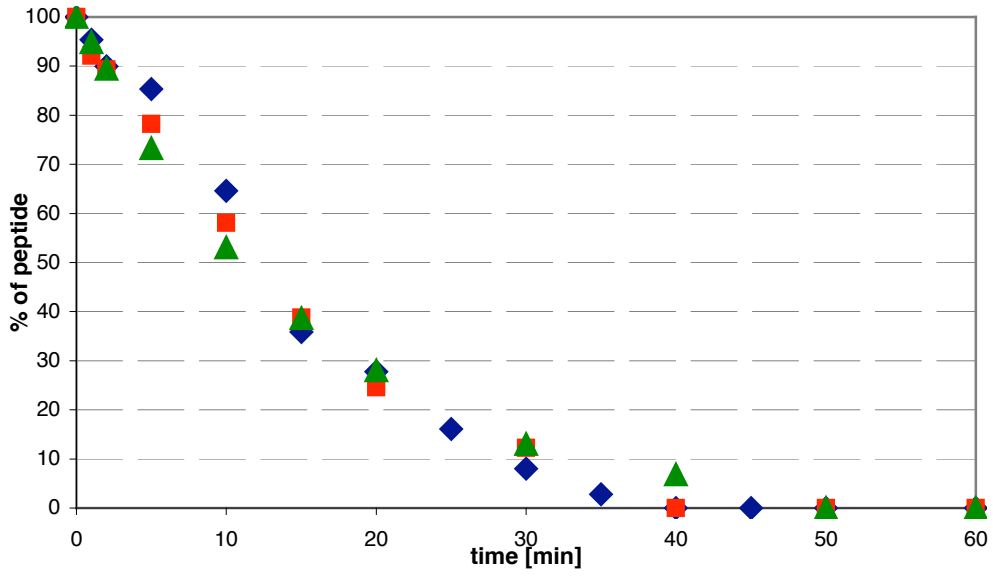



**Figure 48.** The trypsin digestion experiment performed with L4-01. The experiments was repeated in triplicate and resulted in a half-life of 12.5±0.5min.

These results show that the cyclic peptides stabilized in a β-hairpin conformation, are less susceptible to proteolysis than their linear counterparts, and this is consistent with the common assumption that only extended or linear peptide conformations are recognized by trypsin and by other proteolytic enzymes. Presumably the cyclic form is tightly locked into a β-hairpin secondary structure and the peptide bonds are involved in a strong H-bonding network, which has an important influence on recognition by trypsin. For L2-15 and L4-01 the cleavage rate is decreased by around 10-fold, compared to the linear peptide L2-15a, consistent with the β-hairpin-stabilizing effects noted above. In the case of L4-04, which is much more resistant to trypsin proteolysis, there is an interesting correlation with the NMR data of this peptide, which showed very slow H/D exchange rates, consistent with a very stable β-hairpin structure stabilized by strong hydrogen bonding.

#### 2.2.8.2. Plasma stability assay

Compounds with certain functional groups are more susceptible to hydrolysis by plasma enzymes than others. These include esters, amides, lactones, lactams, carbamides, and sulfonamides. The plasma stability assay is designed to focus on these classes of compounds. Arg-containing peptides are potential substrates for the wide variety of trypsin-like proteases that are present in the blood and may have residual activity in the blood. The plasma stability of the Tat peptidomimetics may therefore be limited and this would have significant consequences related to development of antiviral Tat peptidomimetics as drug candidates. Due to species differences, screening of plasma stability in the most relevant animal species provides useful information for drug discovery. This information may help to prioritise compounds for *in vivo* studies and to alert researchers to the potential liabilities of key pharmacophores, so that structural modifications can be made to improve stability.

The stability of selected Tat-derived peptidomimetics was assessed by incubation experiments with the selected peptides (10 μM) in human and mouse plasma (90%), and collecting six data points over the course of 240 min. The experiments were repeated in triplicate. Propantheline was used as a positive control, which has a known degradation profile for both plasma species.[127] The results are shown in Table 29.

**Table 29**. Summary of the results from plasma stability assays. The half-life ($\tau_{1/2}$ [min]) and stability after 4 h (% remaining after 4 h) are presented for each peptide (10 μM in PBS) after incubation with human or mouse plasma (90%).

| Mimetic | $\tau_{1/2}$ [min] in Human Plasma | Stability after 4 h in Human Plasma [%] | $\tau_{1/2}$ [min] in Mouse Plasma | Stability after 4h in Mouse Plasma [%] |
|---|---|---|---|---|
| L1-06 | >240 | 85.3 | 19.1 | 0 |
| L1-07 | >240 | 74 | 79.2 | 1.9 |
| L2-15 | 90 | 24.8 | 178.7 | 42.7 |
| L4-01 | 240 | 50 | 99.1 | 38.4 |
| L4-04 | 104.5 | 30.4 | 152.9 | 45.1 |
| L5-11 | >240 | 55.0 | 60 | 5 |

Although human plasma is of most interest, mouse plasma is also used to provide data relevant to lower mammals, which are often used in preclinical studies. All the peptides are relatively stable in human plasma, with a significant amount remaining after 4 h incubation (Table 29). Mouse plasma degrades the peptides generally faster than does human plasma. Similar observations have been made by Janssen et al. using hK2-selective peptide substrates.[128] The difference between human and mouse plasma may be attributed to the large variability of the enzymes among different species. These results provide an indirect measure of the probable plasma half-life in humans, although *in vivo* plasma clearance will be faster due to clearance by the liver and kidneys.

The trypsin digestion and plasma stability experiments show that cyclization of the peptides and designing sequences that form stable β-hairpin conformations has a major effect on the proteolysis stability of the peptides, which is a very important factor when considering drug discovery.

## 2.3. Discussion

### 2.3.1. BIV Tat peptidomimetics

The previous results on the BIV Tat peptidomimetics described in detail in Section 2.2.1 provided us with the opportunity to use rational methods to optimise the activity of the lead peptide BIV2, thus demonstrating that this class of peptidomimetic inhibitors is both potent and specific. While it has been possible in the past to discover small molecule inhibitors of the Tat-TAR interaction, specificity has been difficult to achieve. In this work, we demonstrated that just a few rounds of optimisation led to the discovery of potent inhibitors of the BIV Tat-TAR interaction; remarkably the inhibitors are strongly selective over a structure as closely related as HIV-1 TAR, demonstrating that the structural rigidity imposed by the D-Pro-L-Pro template provides selectivity by locking the pharmacophore in a defined geometry at the peptide-RNA interface.

The analysis of the large libraries L1 and L2 gave a detailed insight into determinants of affinity and specificity. Although the total number of peptides synthesised and tested is relatively large (~100), it represents a small fraction of the total sequence space potentially available to these peptides. The first step in designing the peptides was to mutate single amino acid residues. Then when the optimal residues for each position were found, the next step was to combine the favourable changes in order to verify if the positive results were additive.

#### 2.3.1.1. Summary of the structure-activity-relationship studies

The guanidinium groups of **Arg1**, **Arg3** and **Arg5** with their positive charges and specific hydrogen bonding abilities are important for tight binding and any mutations at these positions, including lysine and ornithine, abolished binding to RNA. Side chains of these arginines are buried in the RNA major groove, and all three guanidinium groups are located in positions where strong hydrogen bonding with base pairs and/or the RNA backbone can occur. By comparing the complexes of BIV Tat-TAR and BIV2-TAR, there is a clear positional overlap of Arg3 with Arg70, Arg1 with

Arg73 and to a smaller extent Arg5 with Arg77. All three arginine residues in the native BIV Tat peptide were found by mutagenesis studies to be crucial for binding.[129]

The side chain of **Val2** is on the solvent exposed side in the complex and possibly therefore mutations here are well tolerated (Table 5). Val2Thr yields a small improvement of affinity, probably since threonine is similar in size and shape to valine but is polar, which is important for the interaction with water.

**Thr4**, similar to Val2, is also solvent exposed and does not contact the RNA – a variety of mutations were tested and proved to be well tolerated. Just in one case, the Thr4Gln mutation, a small increase of affinity was noticed.

**Gly6** and **Lys7** are situated in the tip of the hairpin in BIV2 and do not contact RNA. The Lys7 side chain is pointing out into solution and the variations introduced at this position are well tolerated. Reversing the order of Lys and Gly has no significance on the affinity towards BIV TAR RNA.

**Arg8, Arg9, Arg11** can only be replaced by other positively charged amino acids (Lys, Orn) to produce peptides that can still bind to TAR RNA. The positive charge appears to play an important role in binding, while the guanidinium group with its specific hydrogen bonding ability seems to be less important. **Arg8** is the most sensitive of all three positions to the changes, possibly because the guanidinium group of Arg8 contacts the base pairs at the tip of the RNA hairpin, however, it is not buried in the structure of the complex.

The side chain of **Ile10** is buried in the complex where it contacts RNA bases. Isoleucine can be exchanged by other hydrophobic residues without much loss of affinity in most of cases, but the side chain of isoleucine still remains as the optimal for hydrophobic contacts at the peptide-RNA interface.

**Ile12** (which replaced Val in the BIV2 sequence at an early stage of optimisation, as it produces consistently a small gain in affinity) points into the major groove of TAR. It could be replaced with polar and even aromatic residues, such as Asn, Thr, or Phe, with only small changes in activity.

Peptidomimetics containing multiple substitutions in the BIV2 sequence, which combine the changes that individually resulted in increased affinity, gave several mutants exhibiting improved affinities towards TAR RNA in comparison to BIV2 (L2-32—L2-36).

**2.3.1.2. Selectivity of the BIV Tat peptidomimetics**

The peptidomimetics in the libraries 1 and 2 were designed to target the Tat-TAR RNA interaction in BIV. Since the lack of specificity of inhibitors of protein-RNA interactions is a well-known problem, these peptides were also assayed against HIV-1 TAR and RRE RNA, in order to evaluate their ability to distinguish between these related RNAs. Remarkably, very large differences in affinity between BIV and HIV TAR were observed when even single amino acid mutations were introduced (Table 5 and Table 6). For example, peptides L2-03 and L2-04 bind to BIV TAR only weakly, while these peptides bind to HIV TAR strongly. Thus, the location of the Arg1 side chain on the hairpin scaffold is an important feature not just of the affinity but also of the selectivity of these peptidomimetics. In contrast, peptides L1-09, L1-11—L1-12 bind to HIV TAR weakly (or not at all) compared to BIV TAR, emphasizing the fact that position 10 can be occupied by different hydrophobic amino acids and retain binding to BIV TAR, although Ile seems to be the optimal residue. On the other hand, Ile was the only hydrophobic residue to retain affinity against HIV TAR; all other hydrophobic residues abrogated binding, while the activity with respect to HIV TAR was retained or even increased with the introduction of polar residues such as Asn and Gln in this position. Finally, significant differences in the binding profile between the two highly similar TAR sites were observed for mutants at the β-turn of the peptide loop (peptides L2-15—L2-20) and for peptides in which multiple mutations were combined, such as L2-11—L2-14 and L2-26—L2-43. For example, swapping Gly for Lys in L2-15 led to a 10-fold increase in the level of HIV TAR binding, mutating Lys to different residues (peptides L2-16—L2-20) abolished binding, but binding to BIV was not affected by any of these changes. Thus, this family of β-hairpin mimetics binds with nanomolar affinity and high selectivity to TAR RNAs and discriminates very effectively between the very closely related HIV and BIV structures.

**2.3.1.3. Summary**

This work demonstrates that peptidomimetic chemistry based on structure-promoting scaffolds yields nanomolar inhibitors of viral protein-RNA interactions. In addition to being as potent as the complete Tat protein they were intended to mimic, these peptides are remarkably selective. Mutations that proved to be most favourable

for BIV TAR were not identical to those that led to optimal activity against HIV TAR, and vice versa; single and multiple substitutions allow the peptidomimetics to distinguish between two RNAs that are closely related in structure and sequence. The knowledge generated from the structural analysis and implemented in just a few cycles of design and testing led to a 100-fold improvement in activity compared to that of the lead molecule BIV2. The insight gained through this work into the origins of the affinity and selectivity of binding of peptidomimetics to BIV TAR has recently been exploited in the development of related mimetics that bind tightly to HIV TAR RNA.

### 2.3.2. HIV-1 Tat peptidomimetics

The ultimate goal of this project was to deliver potent, selective and proteolytically stable peptides that are capable of *in vitro* binding to HIV-1 TAR RNA and inhibiting the HIV-1 Tat protein-TAR RNA interaction, as well as being active *in vivo*, suppressing HIV replication. Studies on BIV Tat-TAR interaction inhibitors performed over the last few years have yielded a deeper understanding of their affinity and selectivity in binding TAR RNA. Furthermore, NMR studies of selected peptidomimetics and the structure of the BIV2-BIV TAR RNA complex allowed the successful targeting of the HIV-1 Tat-TAR interaction.

From the collection of more than 100 peptides designed for the BIV project, it was possible to select some already very potent anti-HIV agents, namely L1-06, L1-07 and L2-15 with affinities of 75 nM-150 nM for HIV TAR RNA. However, further optimisation of the lead compounds (library L3) did not yield any affinity improvement. Therefore, the next step was to extend the peptide sequence in order to reach the region of the RNA loop where the binding of cyclin T1 takes place. Following this approach, a series of peptides with different loop-lengths were produced, containing 15, 16 or 18 amino acids with sequences based on that of BIV2 peptide. In addition to the standard D-Pro-L-Pro template, the majority of peptides were stabilised by means of a disulfide bridge. The results showed that only the 18 amino acid peptides proved to be potent TAR binders, 16-mers were binding but with decreased affinities, while 15-mers failed to interact with the RNA target. Further studies of the longer loop peptides, by mutating single Arg amino acids in the lead compound, L4-04 (Arg→hArg and Arg→Lys), allowed the evaluation of their importance in TAR binding. The analysis of the mutations showed that most of the changes are well tolerated, with the

exception of the Arg15Lys mutation, which exhibits a complete loss of activity. Introducing hArg does not yield any improvement. Evaluation of the importance of the restraints (template and disulfide bridge) indicate that although the dual stabilisation is important for producing stable β-hairpin peptides, the disulfide bridge is not necessary for high binding affinity, whereas removal of the template always brings an increase in $K_d$ value (2-4 fold).

By testing various pairs of amino acids defining β-turn in library L5-01—L5-12, it was once more found that Lys-Gly is the best of the tested versions. This library delivered the strongest binder (L5-11) with a $K_d$ value of 25 nM in a standard EMSA (with excess tRNA) and $K_d$=0.5 nM without excess of tRNA in the EMSA.

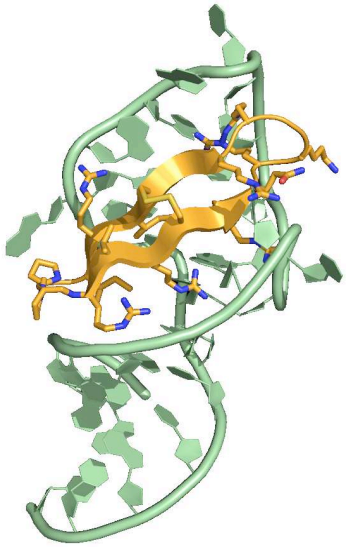


**Figure 49.** Model of complex L5-11-HIV TAR RNA. The NMR-solved structure of L5-11 is superimposed on the HIV-1 TAR RNA NMR structure.

A model of L5-11 complexed to HIV-1 TAR RNA, based on the solved structures of the free peptide (see Figure 49) and of HIV-1-TAR RNA,[72,106] suggest that by extending the peptide loop it might be possible to increase the contact surface between the peptide and the RNA. Also, the interaction would then not only be restricted to the bulge region but would also extended to the RNA-loop region, where it might inhibit the native interaction with cyclin T1.

In parallel to the structure-activity relationship studies and optimisation, selected potent peptides were tested for their stability *in vitro* by means of trypsin digestion assays and plasma stability assays. In comparison to linear peptides, the cyclic structured peptides have at least 10-fold higher stability in the trypsin assay. In

the best case, L4-04, the half-life time was ~360 min. The plasma stability assays, performed in human and mouse plasma, confirmed the former results and proved that the cyclic peptides with a stable β-hairpin conformation are proteolytically more stable, which makes them promising drug candidates.

In the course of these studies, it was also possible to prove that the peptides are cell membrane permeable, which would allow them to perform their inhibitory action in infected cells. Up-take studies performed with fluorescently labelled peptides in HeLa cells demonstrated that the peptides are internalised in a manner similar to the positive control Tat peptides and are localized predominantly in the cytoplasm. Further studies to prove the activity of the peptides *in vivo* and their ability to suppress viral transcription are now being pursued in collaboration with Professor J. Karn, Case Western Reserve University, USA.

Several peptides containing 14 and 18 amino acid residues were studied in detail by NMR techniques. All of them proved to form very stable β-hairpin conformations, as determined using NOE-derived distance restraints. Additional experiments, such as H/D exchange and determination of temperature coefficients for amide bonds suggest that the two antiparallel strands of the β-hairpin are stabilised by hydrogen bonding, especially between amide bonds at positions 1-12 in the 16-mer loop peptides and 3-9 in the 12-mer loop peptides, which exhibit half-life times of exchangeable NHs of several days. The stability of the conformation is of dual importance. First it allows us to establish structure-activity relationships, and second it increases the proteolytic stability of the peptides, which is an important factor when therapeutic use is considered.

# 3. PEPTIDOMIMETICS OF THE HIV-1 REV PROTEIN

## 3.1. Introduction

### 3.1.1. The Rev protein-RRE RNA complex

The virally encoded Rev protein of HIV-1 plays a critical role in viral replication by regulating the transport of unspliced and partially spliced viral RNA from the nucleus to the cytoplasm of infected cells. Rev is a 116 amino acid RNA-binding phosphoprotein that binds a *cis*-acting RNA regulatory element located within the *env* mRNA, termed the Rev response element (RRE). Mutational analysis of Rev protein revealed several discrete domains: **i**) an amino-terminal domain that determines RRE binding and nuclear localization (NLS), **ii**) a multimerisation domain, flanking the RRE binding domain, and **iii**) a carboxy-terminal domain that includes a nuclear export signal (NES) and binding site for cellular proteins, known as the activation domain (Figure 50).[130]

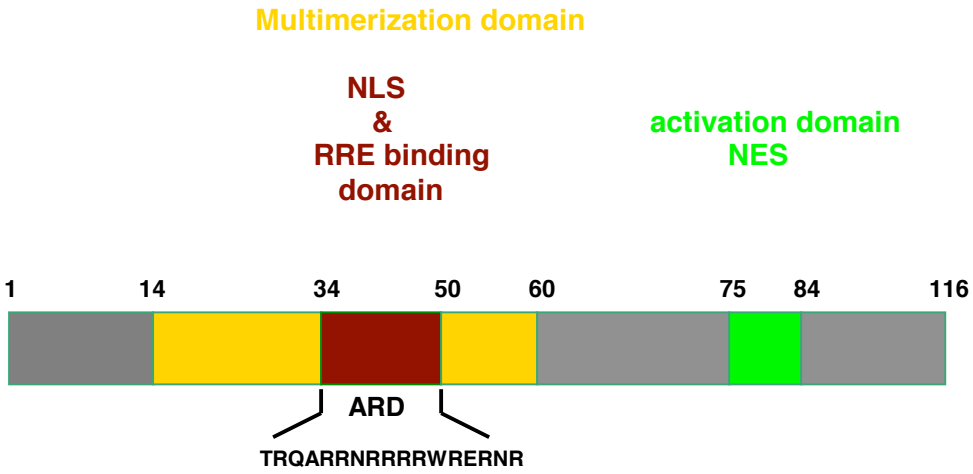



**Figure 50.** The 116 amino acid Rev protein harbours regions that: 1) mediate RRE RNA binding and nuclear localization [Arginine-Rich domain (ARD), brown box], 2) are required for protein multimerization (orange boxes), and 3) which function as the NES/activation domain (green box).[130]

The NLS and RRE RNA-binding domain is rich in basic residues, which makes it, together with the Tat protein, another very prominent example of an RNA-binding protein containing an arginine-rich motif. The target of Rev binding, the RRE, is ca. 350-nucleotides and forms an array of stem-loop structures present within the *env*

intron in all unspliced and singly spliced viral mRNAs. It has been demonstrated by various methods, including mutagenesis and chemical modification interference, that the Rev binding site is located in the 13-nucleotide bulge structure in stem-loop IIB (Figure 51). Although the stem-loop region is helical in nature, the local structure is substantially distorted by the formation of two non-Watson-Crick purine-purine base pairs (G47-A73 and G48-G71).[131] These non-canonical base pairs open the major groove of the A-form RNA double helix, making the bases more accessible to the arginine rich, positively charged Rev peptide.

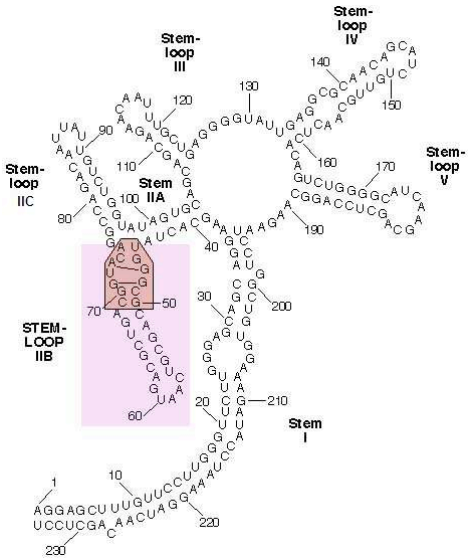



**Figure 51.** Structure of the RRE RNA. The stem-loop IIB (shaded) contains the Rev binding site (in the box). The two non-canonical purine-purine base pairs are indicated.[132]

In 1996 two independent groups succeeded in the determination of the three-dimensional structure of the high affinity RRE site complexed with an arginine rich Rev-derived peptide (Figure 52).[25,133] These NMR studies revealed the α-helical character of the Rev peptide binding in the major groove of the RNA near the purine-rich bulge. The structure formed by the two purine-purine base pairs of the RRE creates a distinctive binding pocket for specific Rev peptide recognition. According to the structure of the complex, arginine side chains at positions 35, 39 and 44 make base-specific contacts, and the side chain of Asp40 contacts a non-Watson-Crick G-A base pair (G47-A73). The side chain protons of Thr34, Arg38, Arg41, Arg42, Arg46, Arg48 and Arg50 are within hydrogen bonding distance of the backbone phosphates. Ala37 is involved in a hydrophobic interaction by stacking of the methyl group with the sugar ring of G46 in the complex. These observations are in agreement with circular dichroism and mutation experiments, which indicate that Thr34, Arg35, Arg38, Arg39,

Asn40 and Arg44 are critical for specific high affinity RNA binding.[134] The α-helical structure of the peptide is necessary for specific-RNA binding and is stabilized when bound to RRE RNA. The RRE RNA, like the TAR RNA (Paragraph 2.1.2) undergoes a conformational change upon binding, and the major groove is widened by ~5Å relative to the unbound form.[135]

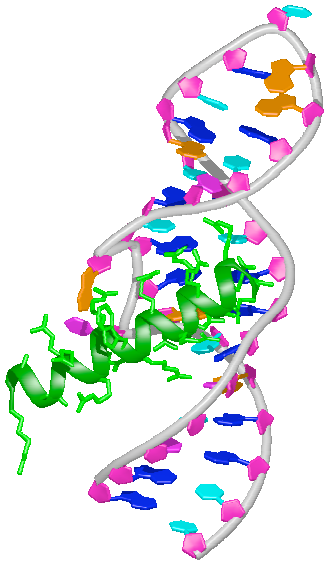


**Figure 52.** Structure of the Rev peptide (green) bound to HIV-1 RRE IIB. Prepared from protein data base file (PDB 1ETF).

Initially, Rev binds to a high-affinity site within the RRE (stem–loop IIB). Subsequent to this initial binding event, approximately 10 additional Rev molecules oligomerize through protein-protein and protein-RNA interactions and coat the entire RRE.[136]

### 3.1.2. Mechanism of action of Rev protein

HIV-1 transcription yields ~2kb, ~4kb, and ~9kb RNAs. The fully spliced ~2kb transcript encodes the regulatory proteins Tat, Rev and Nef, the partially spliced ~4kb RNA encodes Env, Vif, Vpr and Vpu, and the full-length ~9kb transcript, in addition, encodes Gag (including the structural proteins for the capsid, nucleocapsid and matrix proteins) and Pol (encoding viral enzymes). During the early stages of virus expression, which correspond to no or low levels of Rev, only ~2kb transcripts are able to pass through the nuclear membrane and enter the cytoplasm, while unspliced mRNA

remains in the nucleus. Upon production in the cytoplasm, Rev enters the nucleus and activates the nuclear transport of ~9 and ~4 kb RNAs (Figure 53).

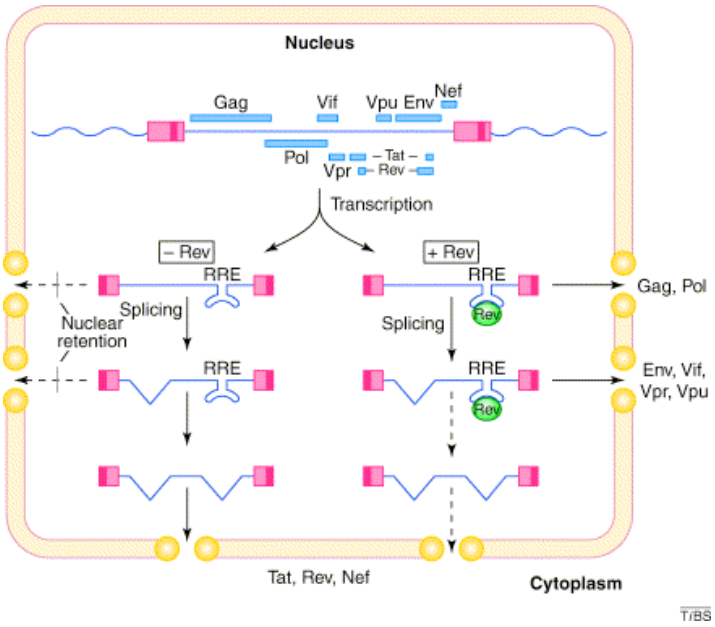



**Figure 53.** Rev is required for cytoplasmic expression of intron-containing HIV-1 mRNA. Schematic representation of early (left) and late (right) phase of mRNA expression.[137]

As the cellular concentration of Rev increases, additional molecules of Rev oligomerize cooperatively along the RRE by binding to adjacent low-affinity sites.[136] By interacting with the nuclear export factor, exportin 1, these multimeric Rev assemblies mediate the export of unspliced or partially spliced HIV-1 mRNAs to the cytoplasm before they can be fully spliced.[130] Since this intron-containing mRNA subset encodes all of the protein components of the mature virus particle, the action of Rev is essential for the production of virions and the progress of HIV-1 infection.[138]

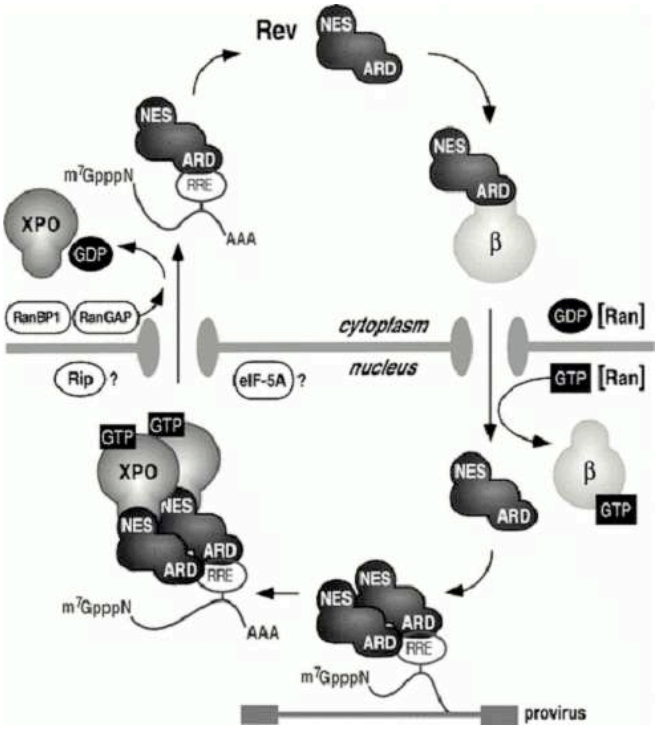



**Figure 54.** The Rev nuclear transport cycle (see text below).[130]

After synthesis, the Rev protein directly interacts with importin β (Imp β) via the nuclear localisation signal domain, located within the arginine-rich domain. The Imp β-mediated nuclear import is known to require energy and involves interaction with specific nucleoporins from the nuclear pore complex (NPC) and the small GTPase Ran. Transport through the nuclear pore complex is driven by a gradient of Ran/GTP, with a high concentration of Ran/GDP in the cytoplasm and a high concentration of Ran/GTP in the nucleus. Once the Imp β-Rev complex reaches the nucleus, the Rev is released from the complex. The next step in the Rev pathway involves the specific assembly of multiple Rev molecules onto RRE and recruitment of exportin-1 (known also as CRM1) via its nuclear export signal domain. The resultant ribonucleoprotein complex is recruited to the nuclear pore complex via direct interaction of exportin-1 with nucleoporins. Once the HIV-1 RNA cargo reaches the cytoplasm, the complex is disassembled and HIV-1 mRNA is available for translation while Rev is free to return to the nucleus (Figure 54).[137,138]

### 3.1.3. Inhibitors of the Rev-RRE RNA interaction

Because of its essential role in viral replication, the interaction between the Rev protein and the high-affinity RRE binding site represents an attractive yet unexploited target for antiviral therapy. Several steps are required for Rev function: binding to the RRE, oligomarization of Rev monomers, and interaction with cellular factors from the nuclear transport machinery. Each of these steps represents a potentially specific target for therapeutic intervention. The fact that structural information for Rev-RRE binding is available makes it a very attractive target for the development of HIV inhibitors.

The first approaches towards anti-Rev drug discovery focused on the Rev-RRE interaction, as it is very specific and has no cellular counterparts. Because Rev binds to an RNA target, intercalating agents with specificity or preference towards RNA were first investigated as potential Rev-RRE inhibitors. The intercalating dye **pyronin Y (28)** was reported to completely block the formation of the Rev-RRE complex *in vitro*, at low μM concentrations. Despite the strong *in vitro* effect, the compound failed to inhibit HIV replication in cytoprotection assays and exhibited high levels of cellular toxicity.[139] Other intercalating agents, such as **derivatives of diphenylfuran (29)**,[140] or **proflavines (30)**[141] were also reported to inhibit the Rev-RRE interaction, by causing a conformational change in RRE. Although these agents can be useful as probes to investigate the precise mechanism of Rev-RRE binding, intercalating agents are clearly not attractive molecules from a therapeutic point of view, because of their toxic and mutagenetic effects.

A second group of compounds shown to bind to RRE *in vitro* and inhibit its interaction with Rev are aminoglycosides. Binding studies with known antibiotics such as **neomycin B (31)** or **tobramycin (32)** showed their ability to bind to RRE at low μM concentrations and inhibit Rev binding to its target RNA.[142] However, the selectivity of the aminoglycosides as RNA-binders is limited, which limits their usefulness as prospective anti-HIV drugs.

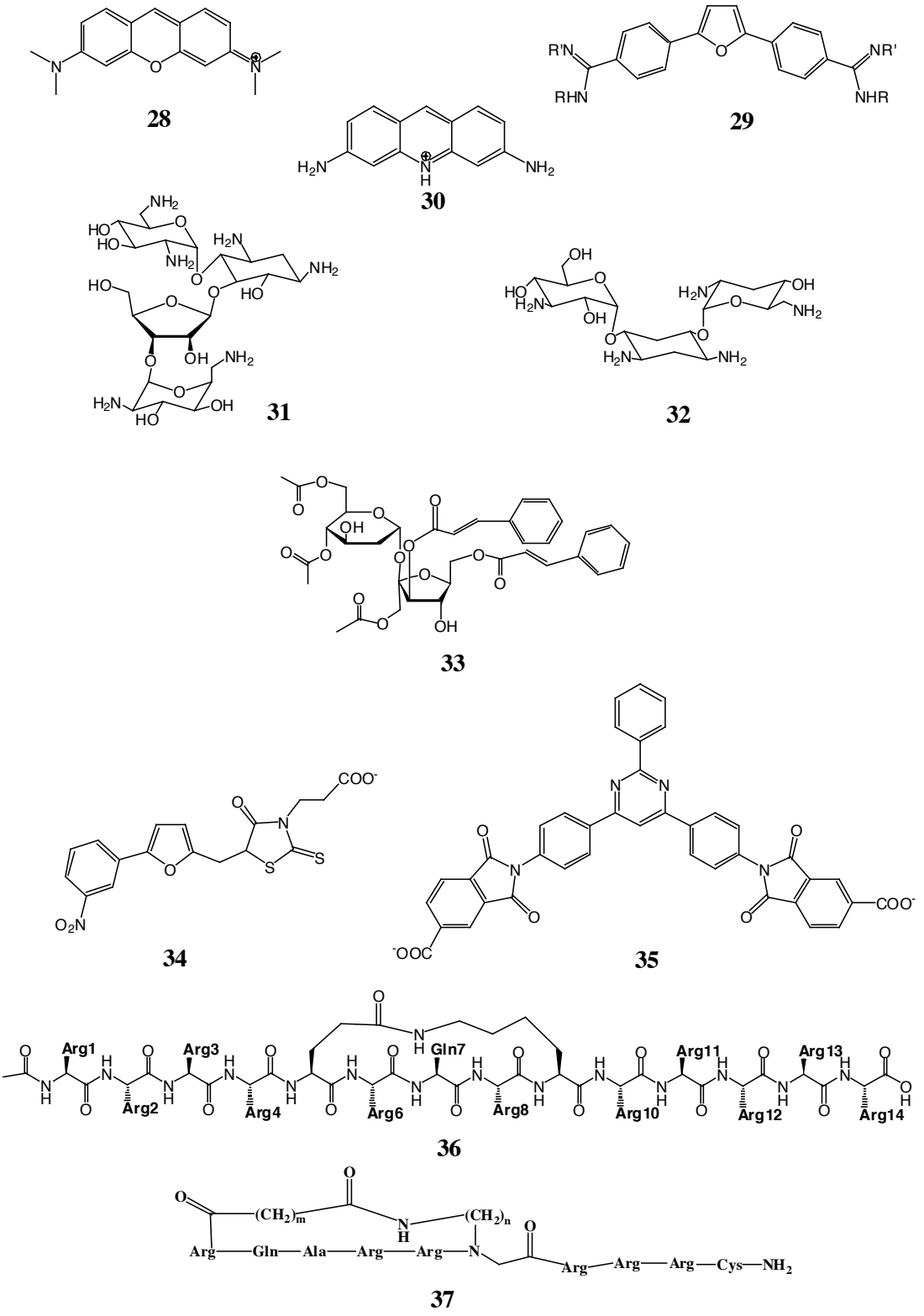



**Figure 55.** Relatively low molecular weight inhibitors of Rev function: (28) pyroninY, (29) diphenylfuran derivatives, (30) proflavine, (31) neomycin B, (32) tobramycin, (33) niruriside, (34) furan derative, (35) complex aromatic system, (36) α-helical constrained peptide, (37) Rev-derived backbone cyclic peptide.

A random screening approach has been utilized to discover compounds capable of inhibiting Rev function. A screening of natural products for their ability to inhibit binding of Rev protein to [$^{32}$P]-labelled RRE RNA, reported by Qian-Cutrone,[143] delivered one novel compound, **niruriside (33)**, which was isolated from the methanol extract of the dried leaf of *Phyllanthus niruri* (a plant widely used in Indian traditional medicine) by bioassay-guided fractionation. This compound was shown to inhibit binding of Rev to RRE with an $IC_{50}$=3 μM. However, niruriside did not protect CEM-SS cells from acute HIV infection.

A high-throughput scintillation proximity assay with biotinylated Rev protein and RRE RNA was used by Chapman et al.[144] to screen over 500'000 small molecules. Several inhibitors were identified, mostly aromatic-rich compounds e.g. **disubstituted furan derivative** (**34)** or a compound containing **complex aromatic system** (**35**). However, no compound demonstrated HIV antiviral activity in a relevant cell-based assay.

The peptidomimetic approach to discover potent Rev-RRE inhibitors was reported by Chaloin[145] and Mills.[146] Mills et al. presented a set of **α-helical arginine-rich constrained peptides**, among which one **(36)** exhibits a high affinity ($K_d$~40 nM) and specificity towards RRE RNA. It was also the most effective competitor ($IC_{50}$~150 nM).

Chaloin reported also a series of backbone cyclic peptides derived from the arginine rich motif of Rev (**37**). All of the peptides exhibit inhibition in the low μM range, and they were able to block HIV-1 replication in cell cultures in early and late stages of virus infection. Reported data suggest that these peptides represent an interesting approach for the design of novel antiviral compounds, however, the concentration required for complete inhibition has to be improved.

All of these approaches are in an early phase of development, and these first attempts represent proof-of-principle experiments indicating that Rev-RRE interactions can be disrupted by small molecules, and that it is a valid molecular target for drug discovery.

### 3.1.4. Project outline

Due to its essential role in the virus life cycle, the viral regulatory protein Rev and its highly specific interaction with the RRE RNA constitutes an attractive target for development of new antiviral molecules. With the experience gained in the field of viral protein-RNA interaction inhibitor (Chapter 2), and being encouraged by recent success in using β-hairpin structures to mimic an α-helical epitope in p53 and inhibit the p53-HDM2 interaction (Figure 56),[147] a new approach to inhibitors of the Rev-RRE interaction was envisioned, based on conformationally constrained β-hairpin peptidomimetics.

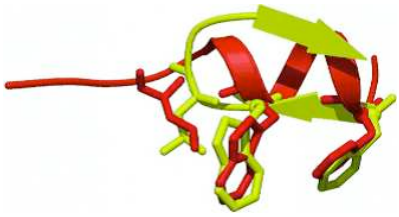


**Figure 56.** Using a β-hairpin to mimic an α-helical epitope. A model β-hairpin (yellow) superimposed on the p53 helical peptide (red).

The RNA-binding sequence of Rev belongs to the same class of protein (arginine-rich domains) as the Tat proteins of BIV and HIV. Having successfully mimicked BIV Tat using macrocyclic ß-hairpin peptidomimetics, it was hypothesised that assaying selected Tat peptidomimetics, which display similarities in sequence with Rev-derived peptide, might provide new leads as Rev-RRE inhibitors. Furthermore, it was envisioned that a β-hairpin might provide a robust scaffold upon which the groups critical for RRE recognition could be displayed, just as the α-helical scaffold of the basic domain of Rev presents the energically important residues to the RNA.

It was, therefore, decided to first assay selected β-hairpin peptides derived from the Tat peptidomimetic collection and, second to design hairpin peptides based on the sequence of the Rev protein, to enable the discovery of potent and specific Rev-RRE interaction inhibitors.

## 3.2. Results

### 3.2.1. BIV Tat peptidomimetics as Rev-RRE inhibitors

Owing to the high similarity of the arginine-rich regions of Tat and Rev proteins, the BIV Tat-derived β-hairpin peptidomimetics (BIV1-BIV8) were tested for *in vitro* binding to RRE, using EMSA. The Rev-derived linear peptide (H-TRQARRNRRRRWRERQR-$NH_2$ (Rev34-50)) was used as a positive control in all binding assays.[148,149] The dissociation constant ($K_d$) of this Rev-derived peptide was determined by EMSA to be approximately 100 nM in the presence of a large excess of *E. coli* tRNA that was used as a control for non-specific binding (Figure 57, left). In the absence of excess tRNA, the $K_d$ of Rev-RRE binding was about 20 nM, which is comparable to previous reports.[149]

**Table 30.** Sequences of β-hairpin peptidomimetics BIV1—BIV8. Positions 1-12 refer to residues 1-12 as mounted on the D-Pro-L-Pro template. The dissociation constants ($K_d$) were determined for the interaction of each mimetic with HIV-1 $RRE^{41-79}$ RNA (EMSA) (nb=no binding detected).

| **Mimetic** | **Position** | | | | | | | | | | | | **$K_d$ [μM]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | |
| **Rev (34-50)** | **H-TRQARRNRRRRWRERQR-$NH_2$** | | | | | | | | | | | | **0.1** |
| **BIV1** | I | R | G | T | R | G | K | R | R | I | R | V | nb |
| **BIV2** | R | V | R | T | R | G | K | R | R | I | R | V | 5 |
| **BIV3** | I | Y | R | T | R | G | K | R | R | I | R | T | nb |
| **BIV4** | Y | R | G | T | R | G | K | R | R | I | Y | V | nb |
| **BIV5** | **R** | **R** | **G** | **T** | **R** | **G** | **K** | **R** | **R** | **I** | **G** | **R** | **0.3** |
| **BIV6** | V | R | G | T | R | G | K | R | R | I | K | Y | nb |
| **BIV7** | V | R | R | T | R | G | K | R | R | I | K | Y | 5 |
| **BIV8** | K | R | G | T | R | G | K | R | R | I | G | Y | nb |

Among the peptidomimetics tested **BIV-5** exhibited the highest affinity for RRE with a $K_d$ of 300 nM in the presence of tRNA, only three-times higher than that measured for the wild-type Rev peptide under the same conditions. Two additional molecules, **BIV-2** and **BIV-7**, were also found to bind to the RRE RNA with $K_d$ values in the low micromolar range. The other mimetics failed to bind even at high micromolar concentrations. It was shown earlier that **BIV-2** was the tightest-binding ligand for BIV TAR ($K_d$=0.15 μM), followed by **BIV-5** ($K_d$=1-2 μM), whereas for **BIV-7** no binding to TAR was detected (Paragraph 1.1.1). **BIV-5**, was also tested in an inhibition assay to measure its ability to displace Rev peptide in a preformed Rev-RRE complex. **BIV-5** was found to displace the Rev peptide from the Rev-RRE complex

with an $IC_{50}$ of about 300 nM (Figure 57, right), which is similar to the estimated $K_d$ for direct binding to the RRE.


**Figure 57.** Left: binding assay (EMSA) for the wild-type Rev peptide to RRE RNA (1 nM); Right: inhibition of Rev peptide (0.1 μM) from a preformed complex with RRE RNA (1 nM) by peptidomimetic BIV5 (all concentrations in μM). The lower band corresponds to free RNA, the upper to bound RNA.

### 3.2.2. Design of Rev peptidomimetics as Rev-RRE inhibitors

- **LIBRARY R1: Initial design efforts**

These promising results prompted the synthesis of a dedicated library of β-hairpin Rev peptidomimetics (**R1-01** to **R1-12**), which were made to test the new design concept illustrated in Figure 58 and Figure 59. In the Rev-RRE complex the key RNA-interacting side chains are displayed around almost the entire circumference of the Rev α-helix. It was envisioned that a β-hairpin motif could be used as a structural platform to organise the amino acid side chains important for recognition of RRE RNA in a geometry similar to that present in Rev (Figure 58).

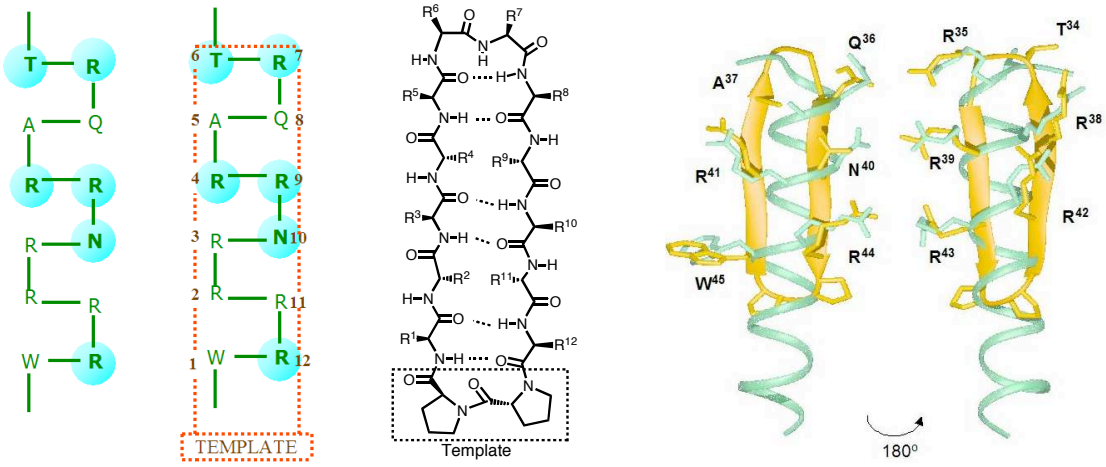



**Figure 58.** Left: Sequence of the α-helical RNA binding domain of Rev protein ($Thr^{34}$-$Trp^{45}$) with the amino acids important for binding highlighted, and a projection of a template-bound cyclic structure with 12 residues overlaid on the Rev α-helix. Middle: a prototypical 2:2 β-hairpin mimetic (residues 1-12 with D-Pro-L-Pro). Right: a model of β-hairpin (yellow) superimposed (the C(β) atoms of the side chains shown were used as the anchor points for the superimposition) on the Rev helical peptide (green). The β-hairpin backbone can be used as a scaffold to pre-organize side chains in a geometry similar to that seen in the helical peptide.

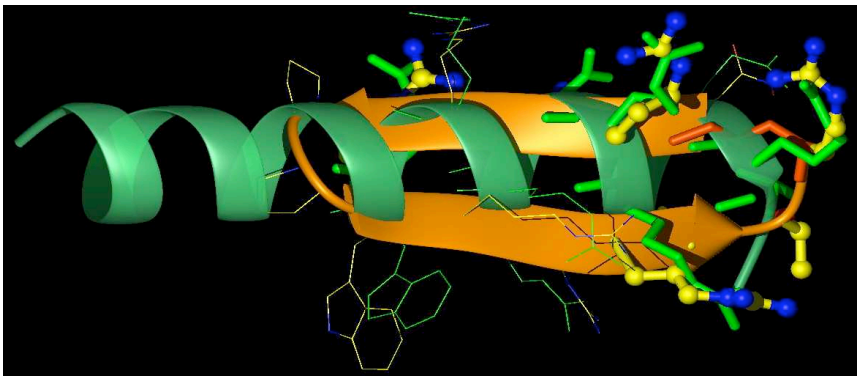


**Figure 59.** A model β-hairpin (orange) superimposed on the Rev-helical peptide (green).

Superimposing a model β-hairpin (Figure 59) on the Rev helical peptide, revealed that the side chains of residues 1, 12, 3, 10, 5, and 8 on one face of the 12-mer hairpin have approximately the same orientations as the side chains of residues W45, R44, R41, N40, A37, and Q36 in Rev, and residues T34, R35, R39, R38, R42, and R43 in the Rev helix might be spatially mimicked by corresponding residues at positions 6, 7, 9, 4, 2, and 11, respectively, on the other face of the hairpin (Figure 58, right). Moreover, the mimicry appears optimal when Arg7 in the hairpin (which should mimic Arg35 in Rev) has the D-configuration. Thus, the first mimetic (R1-01) has the sequence $W^{1}R^{2}R^{3}R^{4}A^{5}T^{6D}R^{7}Q^{8}R^{9}N^{10}R^{11}R^{12}$ mounted upon the D-Pro-L-Pro template. Other mimetics from this library (R1-02 to R1-12) differ in sequence mainly at the tip of the β-hairpin.

**Table 31**. Sequences and binding affinities to RRE determined by EMSA of peptides from the R1 library of Rev peptidomimetics. Residues 1 and 12 are attached to the D-Pro-L-Pro template.

| **Mimetic** | **Position** | | | | | | | | | | | | **$K_d$ [μM]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | |
| **Rev** | **H-TRQARRNRRRRWRERQR-$NH_2$** | | | | | | | | | | | | 0.1 |
| **R1-01** | W | R | R | R | A | T | $^{D}$R | Q | R | N | R | R | 1 |
| **R1-02** | W | R | R | R | A | P | $^{D}$R | Q | R | N | R | R | 1 |
| **R1-03** | W | R | R | R | G | P | $^{D}$R | Q | R | N | R | R | 0.5 |
| **R1-04** | W | R | R | R | V | P | $^{D}$R | Q | R | N | R | R | 1 |
| **R1-05** | W | R | R | R | A | G | $^{D}$R | Q | R | N | R | R | 0.4 |
| **R1-06** | W | R | R | R | A | G | K | Q | R | N | R | R | 0.3 |
| **R1-07** | W | R | R | R | A | T | $^{D}$R | G | R | N | R | R | 0.3 |
| **R1-08** | W | R | R | R | A | P | $^{D}$R | G | R | N | R | R | 0.5 |
| **R1-09** | W | R | R | R | G | P | $^{D}$R | G | R | N | R | R | 0.5 |
| **R1-10** | W | R | R | R | G | T | $^{D}$R | Q | R | N | R | R | 0.7 |
| **R1-11** | W | R | R | R | V | G | $^{D}$R | Q | R | N | R | R | 0.5 |
| **R1-12** | W | R | R | R | A | S | $^{D}$R | Q | R | N | R | R | 0.4 |

The Rev mimetics were synthesized by first assembling a linear peptide chain on 2-chlorotrityl chloride resin using standard Fmoc solid phase peptide synthesis, as described previously in Paragraph 2.2.2. The linear precursors were cleaved from the resin without removing the side chain protecting groups. Macrocyclization was then performed in a dilute DMF solution. The side chain protecting groups were removed with TFA, and the products were purified by reverse phase HPLC. In a typical synthesis, the main component in the crude product was the desired Rev mimetic. After purification by HPLC, the desired peptidomimetics were typically obtained in 20-60% overall yield, and each was >95% pure as determined by analytical HPLC and gave ESI MS data consistent with the calculated masses (see Appendix 3 for analytical data).

The wild-type linear peptide Rev (34-50) was used as a control in all binding studies. The $K_d$ of this peptide determined by EMSA, described previously in Paragraph 2.2.3, in the presence of a large (10'000-fold) excess of tRNA, was found to be approximately 100 nM. The radiolabelled RRE RNA (and TAR RNA for specificity studies) for assays was prepared by *in vitro* transcription using T7 polymerase, synthetic oligonuclotide templates and nucleotides including a portion of [$\alpha$-$^{32}$P]CTP.

All the peptides exhibited a strong affinity to RRE under the conditions tested (Table 31). The model sequence, R1-01, shows a $K_d$=1 μM, which is however lower than the value for BIV5. A single mutation in R1-01 at position 8 (Gln8Gly), as in peptide R1-07, improves the affinity 3-fold to a $K_d$ ≈300 nM. The 4 different pairs of amino acids in the turn region were tested (Thr-$^{D}$Arg, Pro-$^{D}$Arg, Gly-$^{D}$Arg, Gly-Lys, Ser-$^{D}$Arg) and they all were equally well tolerated. However, NMR spectroscopic studies on several of these mimetics in free aqueous solution, suggested that few regular β-hairpin structures were present, that is, that they are disordered. For example, $^{3}J$(NH,C($\alpha$)H) values were all in the range 6-8 Hz, and few characteristic long-range NOEs could be detected. Also, both *cis* and *trans* peptide bond rotamers at the Arg12-D-Pro13 peptide bond can be detected in slow exchange in the NMR spectra. In the case of R1-06, at least one *cis* isomer (at the Arg12-D-Pro13) in equilibrium with the major *trans* isomer was observed, in an approximate 15:85 ratio. However, the peptide backbone NH and H-C($\alpha$) resonances for the major major *trans* form could be assigned unambiguously (see Appendix 6), which enabled calculations of average solution structures calculations for the major isomer using the available restraints (for summary see Table 32). The resulting structures, shown in Figure 60, indicate that the loop of the

mimetic is disordered in free solution. Although the binding assays show high-affinity binding to the RNA, the lack of stable hairpin structure in these free ligands in solution prompted further studies aimed at stabilizing the hairpin fold.

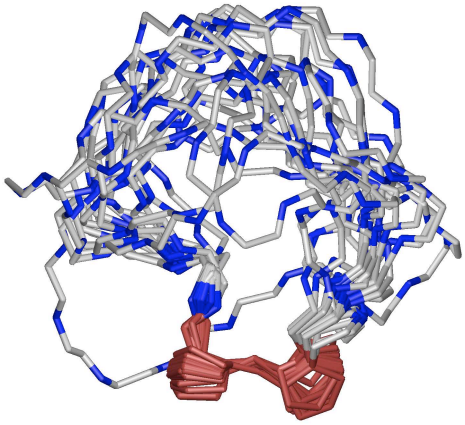


**Figure 60**. Example of a poorly structured Rev peptidomimetic from the R1 library. Superposition of the template and representation of the final 20 NMR structures of R1-06. The structure calculation were performed only on the *trans* Arg12-D-Pro13 peptide bond isomer (for details see text).

| **Mimetic** | **R1-06** |
|---|---|
| **NOE upper-distance limits** | 110 |
| Intraresidue | 52 |
| Sequential | 42 |
| Medium and Long-Range | 16 |
| **Residual Target Function Value (Å$^2$)** | 0.59±0.09 |
| **Mean rmsd values (Å)** | |
| All backbone atoms | 2.21±0.71 |
| All heavy atoms | 4.52±1.00 |
| **Residual NOE violations** | |
| Violation > 0.22 Å | 8 |
| Maximum violation (Å) | 0.37 |

**Table 32.** Experimental distance restraints and statistics for the final 20 NMR structures calculated for R1-06.

- **LIBRARY R2: Attempts to improve the β-hairpin stability of the peptidomimetics**

The second group of mimetics tested (**R2-01** to **R2-16**) combine features of **BIV-5** and peptides **R1-01** to **R1-12**, and include a $K^6G^7$ motif at the hairpin tip, which was expected to promote a stable β-turn, as seen in closely related TAR-binding

mimetics. In addition, a disulfide bridge is present at a non-hydrogen-bonding position in **R2-14**, to further stabilize a β-hairpin structure.

**Table 33.** Sequences and binding affinities to RRE determined by EMSA of peptides from the R2 library of Rev peptidomimetics. Residues 1 and 12 are attached to the D-Pro-L-Pro template.

| **Mimetic** | **Position** | | | | | | | | | | | | **$K_d$ [μM]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | ***2*** | ***3*** | ***4*** | ***5*** | ***6*** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | |
| **Rev(34-50)** | **H-TRQARRNRRRRWRERQR-$NH_2$** | | | | | | | | | | | | 0.1 |
| **BIV5** | **R** | **R** | **G** | **T** | **R** | **G** | **K** | **R** | **R** | **I** | **G** | **R** | **0.3** |
| **R1-01** | **W** | **R** | **R** | **R** | **A** | **T** | **$^{D}$R** | **Q** | **R** | **N** | **R** | **R** | **1** |
| **R2-01** | W | R | R | R | A | K | G | Q | R | N | R | R | 0.2 |
| **R2-02** | W | R | L | R | A | K | G | Q | R | N | R | R | nd |
| **R2-03** | W | R | Q | R | A | K | G | Q | R | N | R | R | nd |
| **R2-04** | W | R | R | R | A | K | G | Q | R | N | R | V | 0.5 |
| **R2-05** | W | R | L | R | A | $^{D}$R | K | G | R | N | R | R | 0.4 |
| **R2-06** | W | R | L | R | A | K | K | G | R | N | R | R | 0.3 |
| **R2-07** | W | R | K | R | A | G | K | Q | R | N | R | R | 0.4 |
| **R2-08** | W | R | K | R | A | K | G | Q | R | N | R | R | 0.2 |
| **R2-09** | W | R | R | R | A | $^{D}$R | G | Q | R | N | R | R | 0.5 |
| **R2-10** | W | R | L | R | A | $^{D}$R | G | K | R | N | R | R | 0.25 |
| **R2-11** | R | R | L | R | A | $^{D}$R | G | Q | R | N | R | R | 0.1 |
| **R2-12** | R | R | L | R | A | K | G | Q | R | N | R | R | 0.1 |
| **R2-13** | R | V | R | R | A | K | G | Q | R | N | R | R | 0.3 |
| **R2-14** | R | C | R | R | A | K | G | Q | R | R | C | R | 0.1 |
| **R2-15** | K | R | Q | R | T | K | G | R | R | L | O | R | 0.1 |
| **R2-16** | W | R | R | R | A | G | R | Q | R | N | R | R | 0.15 |

These peptides were prepared using methods described above. The intramolecular disulfide bridge present in R2-14, was formed by oxidation of the pre-purified cyclic peptide containing cysteine residues protected with Acm groups, as previously described in detail in paragraph 2.2.4.1. After purification by HPLC, the desired peptidomimetics were typically obtained in 20-40% overall yield, and each was >95% pure as determined by analytical HPLC and gave ESI-MS data consistent with the calculated masses (see Appendix 3 for analytical data).

The EMSA results for library R2 are summarised in Table 33. All the peptides, except **R2-02** and **R2-03**, exhibit high affinity towards RRE RNA and many of them are more potent ligands than the peptidomimetics from the previous library. Generally incorporating the Lys6-Gly7 in the turn region brings an improvement in affinity, however, when combined with a mutation at position 3 (Arg3Leu, Arg3Gln), which reduces the total charge of the peptide, a complete loss in activity was observed. It is interesting to notice that these mutations are tolerated if the charge of the peptide is maintained by replacement of Trp by Arg in position 1, as in peptide **R2-12**.

Tryptophan at position 1 is not critical for the interaction with the RNA, it can be replaced by Arg or Lys, as shown in peptides **R2-12—R2-15**. Introduction of a disulfide bridge at a non-hydrogen-bonding position in peptide **R2-14**, apart from being an important stabilisation factor, produces a highly potent peptide. This together with peptides **R2-11**, **R2-12**, and **R2-15** proved to be the best ligands for RRE RNA with $K_d \cong 0.1$ μM, similar to the Rev-derived peptide. When these molecules were assayed in the absence of excess tRNA, a $K_d$ of 10 nM was observed for **R2-11** (Table 34) and **R2-12**, whereas **R2-14** showed a $K_d$ of 1-2 nM. The affinity of this latter peptide for the RNA is thus significantly stronger than that of the Rev-derived α-helical peptide under these conditions.

### 3.2.3. Specificity of the Rev peptidomimetics

A further key issue in RNA ligand design is the specificity of RNA binding, which could clearly be influenced by the flexibility of the ligand as well as by the overall positive charge of this class of peptides. As a stringent test of binding specificity to RRE, binding of the mimetics to the structurally very similar HIV TAR RNA was also evaluated. As a control, binding of a peptide containing 7 consecutive Arg residues (Ac-$Arg_7$-$NH_2$) was also tested, which is expected to interact strongly but non-specifically with both RNAs, mainly through electrostatic interactions.

**Table 34.** Binding affinities (in μM) of selected peptides to HIV RRE and TAR RNAs. The $K_d$ values were determined by EMSA in the **absence of tRNA**.

| **Mimetic** | **$K_d$ (RRE)** | **$K_d$ (TAR)** |
|---|---|---|
| **R2-11** | 0.010 | 0.005 |
| **R2-12** | 0.010 | 0.010 |
| **R2-13** | 0.005 | 0.010 |
| **R2-14** | 0.002 | 0.100 |
| **R2-15** | 0.002 | 0.010 |
| **Ac-$(Arg)_7$-$NH_2$** | 0.002 | 0.002 |

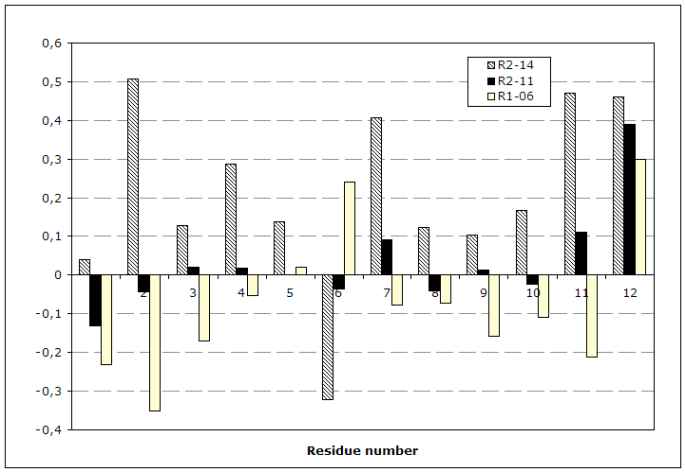
Consistent with this hypothesis, and demonstrated by EMSA, the $Arg_7$ peptide binds non-specifically with a $K_d$ of approximately 2 nM to both HIV TAR and RRE (Table 34). Among the ligands tested **R2-13** and **R2-15** discriminate only modestly (less than tenfold) between RRE and TAR. However, **R2-14** discriminates by approximately 50-fold between the two closely related RNA structures. Given how closely related these RNA targets are, and the relatively poor specificity typically seen

in the interactions of other RNA-binding molecules with more diverse RNA targets, this result represents a notable level of RNA-binding specificity.

### 3.2.4. NMR studies on the Rev peptidomimetics

From the two libraries several peptidomimetics were selected for 2D $^{1}$H NMR experiments in aqueous solution (R1-01, R1-05, R1-06, R1-07, R2-11, R2-14 and R2-16). Except for R2-14, all spectra showed at least two forms in solution with a percentage of minor species between 15-25%. As described previously[54] for similar classes of cyclic peptidomimetics, the two forms are *cis-trans* isomers at the Xxx12-$^{D}$Pro13 peptide bond with the *trans* form the major isomer. Except in the case of R2-14, the NMR spectra revealed predominantly random coiled non-β-hairpin structures in solution, as confirmed by a low amide proton chemical shift dispersion (0.6-0.8ppm), weak up-field shifts of CαH chemical shifts with respect to random coil values (Figure 61), and the absence of characteristic long range NOEs.

**Figure 61.** Hα chemical shifts relative to random coil chemical shifts of residues in peptidomimetics R1-06 (yellow), R2-11 (black), R2-14 (shaded).




In R2-14, only the *trans* form is observed. The formation of the *cis* form appears to be prevented by the additional disulfide bridge restraint. The peptide backbone NH and H-C(α) resonances for R2-14 could be assigned unambiguously (Appendix 6). Several long-range NOEs were observed, indicative of stable structure, which allowed the determination of average solution structures (Table 35). NMR spectroscopic studies on R2-14 in aqueous solution revealed a highly populated β-hairpin conformation in

solution (Figure 62) in which the two strands of antiparallel β-sheet are connected by a type I'-turn between Lys6 and Gly7. In particular, cross-strand NOEs characteristic of a β-hairpin are now observed, and the $^{3}J(NH, C(\alpha)H)$ values are predominantly greater then 8Hz for residues within the β-strands. Thus the disulfide cross-link is clearly important in stabilizing the regular β-hairpin structure.

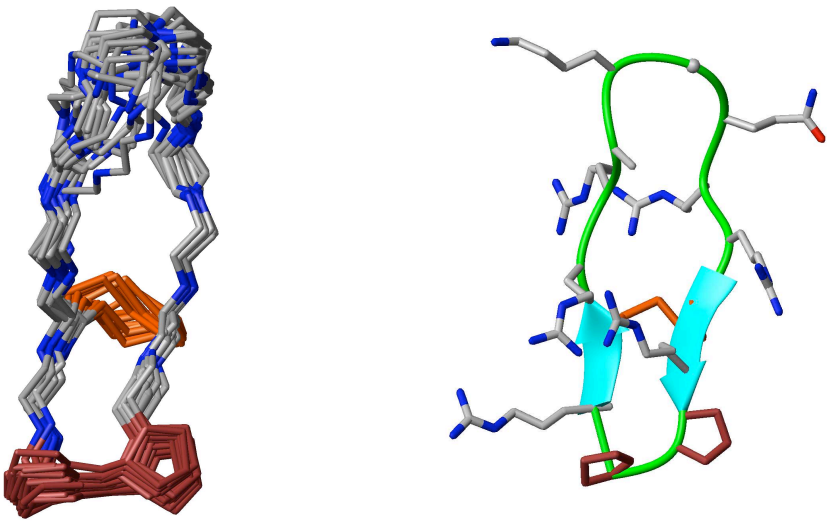


**Figure 62.** Left: Backbone superposition and representation of the final 20 NMR structures. Right: typical NMR structure with the ribbon representation of two strands for **R2-14**.

**Table 35**. Experimental distance restraints and statistics for the final 20 NMR structures calculated for R2-14.

| **Mimetic** | **R2-14** |
|---|---|
| **NOE upper-distance limits** | 109 |
| Intraresidue | 44 |
| Sequential | 50 |
| Medium and Long-Range | 15 |
| **Residual Target Function Value (Å$^2$)** | 0.55±0.03 |
| **Mean rmsd values (Å)** | |
| All backbone atoms | 1.02±0.37 |
| All heavy atoms | 2.57±0.61 |
| **Residual NOE violations** | |
| Violation > 0.22 Å | 2 |
| Maximum violation (Å) | 0.30 |

## 3.3. Discussion

Cyclic peptides containing a D-Pro-L-Pro template are an interesting class of protein mimetics due to the strong β-hairpin inducing properties of the heterochiral diproline. Here we have used this methodology to produce cyclic constrained peptides with the potential to be used as inhibitors of the binding of HIV-1 Rev protein to the RRE RNA. The RNA-binding domain of Rev protein adopts an α-helical conformation upon binding to its cognate RNA. We have used here a novel approach involving use of a model β-hairpin peptide as a scaffold to pre-organise side chains into a geometry similar to that seen in a helical peptide. As a starting point, already available cyclic, β-hairpin peptides derived from the related, arginine rich, RNA-binding sequence of Tat protein were assayed. Identifying one peptide, **BIV-5**, able to bind to RRE RNA with $K_d$=0.3 μM and inhibit the Rev-RRE interaction ($IC_{50}$=300 nM) encouraged further studies on this system. Transposing the RNA-binding domain displayed on a helical segment of Rev protein onto a hairpin scaffold, a potential Rev-RRE interaction inhibitor was designed (**R1-01**). The binding studies showed that this molecule binds to the RNA with $K_d$=1 μM. Further modifications of the sequence, specially in the turn region, led to the first library **R1**, and to the discovery of RRE-ligands with higher affinities ($K_d \approx$ 300 nM). However, the NMR studies on these peptides in free aqueous solution showed that they are largely unstructured, despite the influence of the D-Pro-L-Pro template.

In the second library, **R2**, the interest focused on stabilizing the β-hairpin conformation of the Rev peptidomimetics without loosing the activity of the peptides. For this purpose, information on the sequence of **BIV-5** and the most potent peptides from the **R1**-library was combined. Furthermore, it was envisaged that using Lys6-Gly7 in the turn and incorporating the disulfide bridge would help to stabilize a regular hairpin structure in solution. Most of the peptides from the second library exhibit higher affinities towards RRE RNA (100-500 nM) in standard EMSA than the ones from the previous library. However, the stabilization of the hairpin conformation proved to be a more difficult task than first thought. In the selection of the peptides studied by NMR techniques, only **R2-14**, stabilized by a disulfide bridge in addition to the template, showed a relatively stable β-hairpin structure in free solution. This peptidomimetic exhibits the highest affinity to RRE in the presence ($K_d$=100 nM), as well as in the

absence, of competitor tRNA in EMSA ($K_d$=2 nM). In addition, this compound exhibits good specificity for its target RNA.

The structural and binding data together provide strong evidence supporting the hypothesis that this family of peptidomimetics can mimic the α-helical Rev peptide in its binding to RRE RNA. Although it is not yet possible to rationalize the quantitative effects of individual side-chain substitutions on the binding energy, it was possible to transplant the molecular interactions observed on the α-helical Rev-derived peptide onto a β-hairpin scaffold and rapidly discover potent RRE ligands. Furthermore, the results suggest that not just the number of positively charged groups, but also their relative orientations, as determined by the flexibility of the β-hairpin mimetic, have a decisive influence on the affinity and specificity of RNA binding.

# 4. RNASE P

## 4.1. Introduction

Ribonuclease P (RNase P) is an essential endonuclease involved in processing of precursor tRNA (ptRNA) into its mature form in all three kingdoms of life (Bacteria, Archaea, Eukarya). It has become the focus in the last decade of extensive interest as one of only two known ribozymes conserved in all life domains (the second one is the ribosome) that catalyze reactions with multiple turnovers. In contrast to other ribozymes (like the group I and II introns and the hammerhead RNA), which catalyse intramolecular phosphodiester bond cleavage, RNase P was the first *trans* acting ribozyme identified (i.e. substrate and catalyst are two separate RNA molecules). Although RNaseP catalyses the cleavage of multiple substrate RNAs in the cell, the preferred substrates are tRNAs precursors.[150-152]

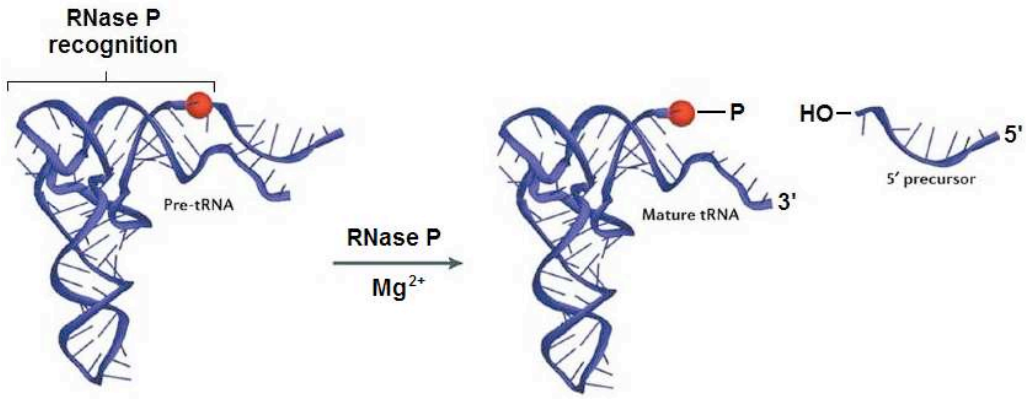



**Figure 63.** The RNase P reaction; processing of precursor tRNA.[153]

RNase P is responsible for the endonucleolytic cleavage of 5'-leader sequences from all precursor tRNAs, producing a mature 5'-phosphate on the end of the tRNA and a hydroxyl group on the 3'-end of the leader sequence (Figure 63). RNase P consists of one RNA subunit and a variable number of protein subunits: only one is present in bacteria, but four are found in archea, nine in yeast and ten in mammalian RNase P.[153]

The RNA component of bacterial RNase P can catalyse ptRNA cleavage in the absence of the protein subunit *in vitro* under conditions of high salt. The RNA subunit consists of separate domains determining specificity and catalysis. However, the protein

subunit is necessary *in vivo* and for the efficient reaction *in vitro* under physiological conditions of low ionic strength.[154-156] The archeal and eucaryal holoenzymes are enriched in protein (70% of total mass of holoenzyme), relative to the bacterial enzyme (10%), which suggest that the protein subunits of these ribozymes perform more complex functions than their bacterial counterparts. It has been suggested that through the course of evolution, the proteins in eucarya and archea RNase P have evolved to play more fundamental roles and their functions might be to stabilize the active conformations of the RNA and to support substrate binding.[150] Only very recently has it been shown that eucaryotic RNase P RNA can also mediate cleavage of a model hairpin-loop substrate as well as tRNA precursors in the absence of the protein subunits, however, the reaction rates and the affinities for the substrate are dramatically reduced.[157] Since the protein component of RNase P is required to activate the ribozyme *in vivo*, it has been suggested that the primary function of the protein is to stabilize the active tertiary structure of RNase P RNA and contribute to high affinity substrate binding.[158]

RNase activity absolutely requires divalent metal ions (preferably $Mg^{2+}$), which probably fill both structural and functional roles in the catalytic cycle (Figure 64).

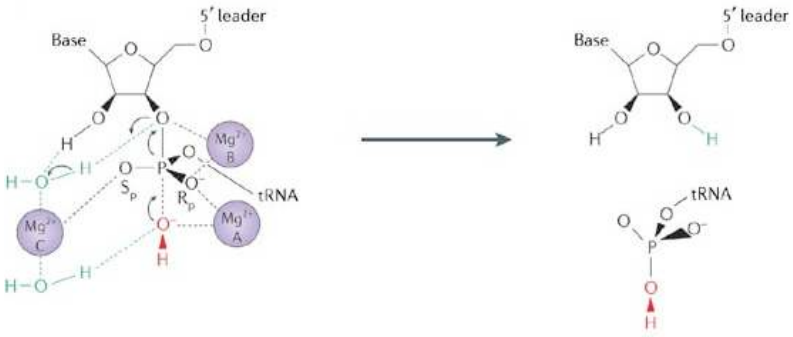



**Figure 64.** The proposed mechanism of the RNase P reaction based on biochemical studies. On the left is a putative structure of the transition state of the reaction. The labile phosphodiester bond is encaged by $Mg^{2+}$-hydrate complexes to form a trigonal bipyramidal transition state. The available biochemical data are consistent with a concerted $S_N2$-like mechanism.[159]

The comparative analysis of RNA sequences has resulted in refined secondary structure models of the RNase P RNA's for the bacterial type A and type B, as well as archeal RNA and eucaryal RNA. Archeal RNase P RNA tends to be similar in sequence and structure to the bacterial one and usually contains extra helices in comparison to eucaryal RNA. RNA sequences from eucaryal RNase P's are simplified (~30% shorter)

in comparison to the other 2 types and lack the key stabilizing helices found in bacterial RNAs (Figure 65).[160]

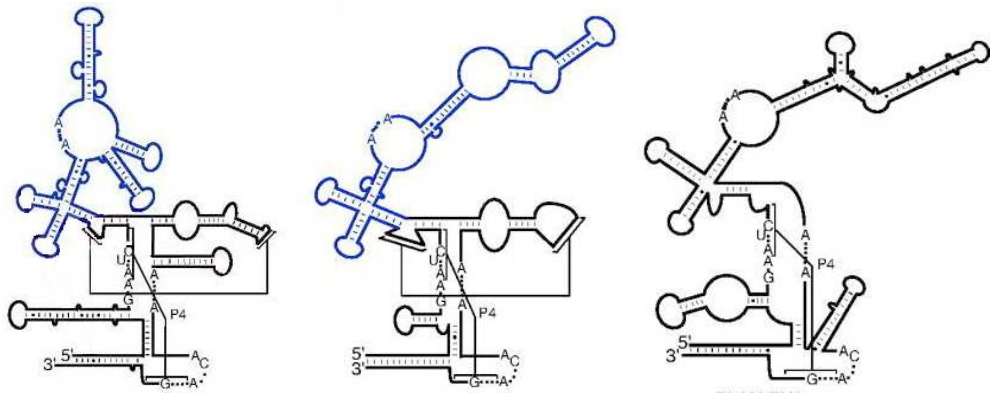



**Figure 65.** Representative secondary structures of RNase P RNA from all three domains of life, bacteria, type A (*Eschericha coli*, also termed M1 RNA), archaea (*Pyrococcus furiosus*), and eucarya (*Homo sapiens*). The blue area in the bacterial and archeal RNase P represents the substrate-specificity domain. Helices P4 and P6, which are disrupted in this representation are connected by line segments. Universally conserved nucleotides are depicted in the presented structures.[160-162]

### 4.1.1. Bacterial RNase P

Bacterial RNase P was one of the first catalytic ribozymes identified and thoroughly studied. The bacterial RNase P RNA component consists of ~300-450 nucleotides and can be subdivided in two major types, A (ascestral-type, such as *Eschericha coli*) and B (e.g. *Bacillus subtilis*), on the basis of their sequence characteristics (Figure 66).

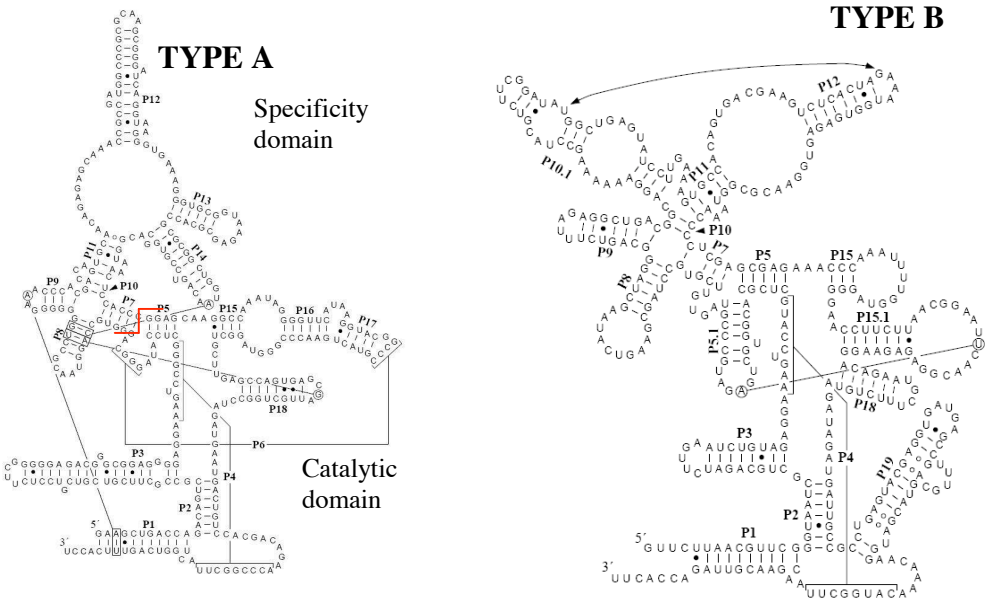



**Figure 66.** Secondary structures of two types of bacterial RNase P RNAs: left *E. coli* (type A), right *B. subtilis* (type B).[163] The red line separates the specificity domain and the catalytic domain.

Since 2003 crystal structures have been available for one of the structural domains of the RNA component (the specificity domain) of RNase P from two bacteria, *Bacillus subtilis*[164] and *Thermus thermophilus*,[165] and more recently for the entire RNA in RNase P from *T. maritima*[166] and *B. stearothermophilus*.[167]

The ternary structure of the protein subunit of *B. subtilis*, *Staphylococcus aureus* and *T. maritima* (representing both A- and B-types of RNase P) have been determined by X-ray crystallography and NMR.[168-170] Although the sequences of these proteins are not highly conserved (20-30% amino acid identity), the overall three-dimensional structures are remarkably similar. The structural similarity results in functional similarity: RNase P RNA and proteins from different organisms can be reconstituted in vitro to form active heterologous complexes.[155,171]

Three regions of the protein have been proposed to interact with RNase P RNA, a basic domain containing an RNR motif, the central cleft formed by one α-helix (C) and four β-sheets, and the negatively charged metal binding loop (Figure 67).[168]

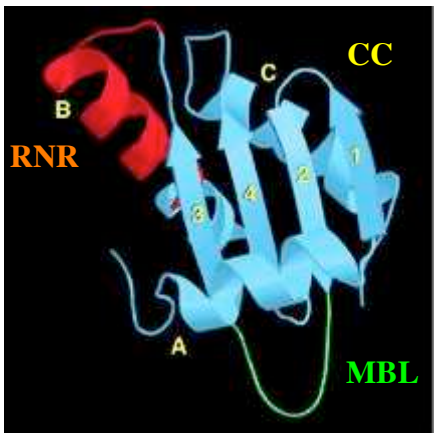



**Figure 67.** Ribbon plot of RNase P protein from *B. subtilis*;[168] α-helices (A=α1, B=α2, C=α3) and β-strands labelled sequentially by letters and numbers, respectively. The basic amino acid-rich conserved RNA binding motif (RNR motif) is shown in red, a negatively charged metal binding loop is indicated as MBL, and the central cleft is indicated as CC.

The RNR motif is a highly basic region located in an unusual left-handed βαβ crossover connection from β-strand 3 to α-helix B to β-strand 4, and it is highly conserved among all bacterial RNase P proteins (Figure 68).

| | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 52 | | 54 | | | | | 59 | | | 62 | | | | | 66 | | | | 70 | | | | | |
| *E. coli* | V | A | K | K | N | V | R | R | A | H | E | R | N | R | I | K | R | L | T | R | E | S | F | R | L | R |
| *C. burnetti* | A | S | K | R | N | V | R | K | A | V | W | R | N | R | V | R | R | V | V | K | E | A | F | R | I | R |
| *N. gonorrhoeae* | V | G | K | K | T | A | K | R | A | N | E | R | N | Y | M | K | R | V | I | R | D | W | F | R | L | N |
| *S. bikiniensis* | V | - | S | K | A | V | G | G | A | V | V | R | N | Q | V | K | R | R | L | R | H | L | V | C | D | R |
| *B. subtilis* | V | - | S | K | K | I | G | N | A | V | M | R | N | R | I | K | R | L | I | R | Q | A | F | L | E | E |
| *Synechocystis* | V | S | Q | K | V | S | K | K | A | T | V | R | N | R | L | K | R | Q | I | R | A | V | I | N | H | F |
| *B. burgdorferi* | F | - | S | K | G | F | R | G | S | V | K | R | N | R | I | R | R | L | F | K | E | A | F | R | K | R |
| *C. pneumoniae* | V | - | S | K | K | F | G | K | A | H | E | R | N | S | F | K | R | V | V | R | E | V | F | R | H | V |
| *T. maritima* | V | - | K | R | K | F | G | K | A | T | R | R | N | K | L | K | R | W | V | R | E | I | F | R | R | N |
| *P. gingivalis* | V | A | K | K | R | F | R | R | A | V | K | R | N | R | V | K | R | L | V | R | E | A | Y | R | L | N |
| *D. radiodurans* | V | S | K | K | T | L | K | H | A | V | K | R | N | R | A | R | R | R | V | R | E | A | L | R | T | M |
| *C. tepidum* | V | G | K | K | L | V | P | R | A | V | D | R | N | R | I | K | R | L | M | R | E | A | Y | R | L | E |

**Figure 68**. Alignment of the amino acid sequence around helix B (α2) in the protein subunit of RNase P from 12 different bacteria.[172] Amino acids that are highly conserved are shown in red.

Among all RNase P, the ones from bacteria, specially from *E. coli* (where the ribonuclease was first discovered[173,174]) and *B. subtilis* (the X-ray structures of the specificity domain of the RNA and the protein subunit are available) are the best characterised. As they consist of a single RNA and a single protein subunit they provide a simple model system to study the function of the protein in the ribonucleoprotein (RNP) complex. Although the structure of the *E. coli* RNase P protein subunit (commonly named as C5 protein) has so far not been solved, it is thought to be similar to the RNase P protein of *B. subtilis*, since the proteins from these bacterial RNase P can be functionally exchanged.[171] Furthermore, there is a common subset of highly conserved basic and hydrophobic residues located in the central region, including the RNR motif, as well as a number of conserved hydrophobic amino acids in the N-terminal and C-terminal domains (Figure 69).

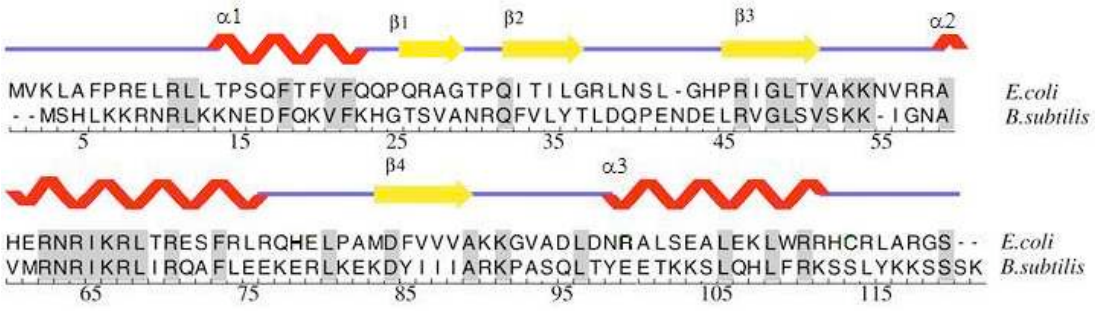



**Figure 69.** A comparison of the sequences of the protein of RNase P from *E. coli* and *B. subtilis*. The identical residues are highlighted in grey. The secondary structure elements are indicated above the sequences.[175]

Although the ternary structures of RNA and protein subunits have been determined, a deeper understanding of the RNase mechanism and function requires detailed structural information of the entire holoenzyme, which is so far lacking. However, mutagenesis, photo-affinity cross-linking and nucleotide analogue

interference mapping (NAIM) studies have identified **nucleotides** essential for ptRNA binding and catalysis.[152] Similar experiments were performed to understand the **protein** subunit function and to identify the regions that bind to either RNase P RNA or to the substrate. Thus cross-linking, affinity cleavage experiments, footprinting experiments, and mapping RNA-protein interactions using disulfide-linked EDTA-Fe have begun to reveal potential RNA-binding domains.[176-178] Combining all the structural information available including the X-ray structural data and the results of various cross-linking techniques, has led in the last few years to many models of the holoenzyme alone or complexed with ptRNA (Figure 70), as suggested by Sharkady and Nolan,[175] Rox at al.,[179], Tsai et al.,[180] Buck et al.,[156] and Niranjanakumari et al.[181]

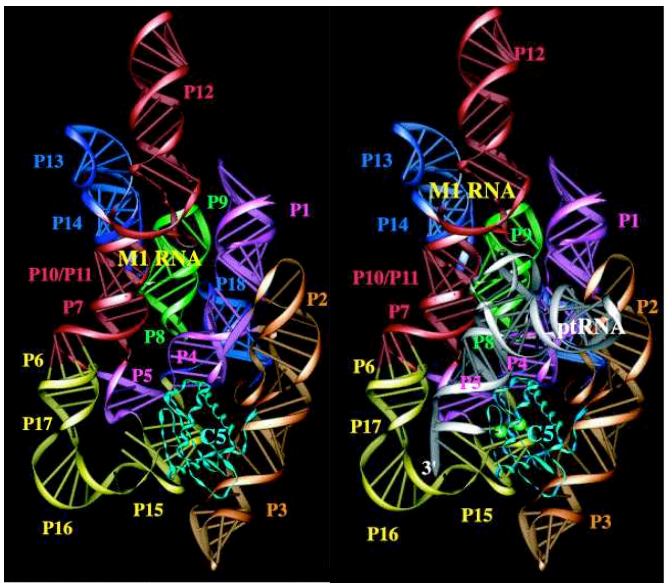



**Figure 70.** Three-dimensional models of *E. coli* RNase P holoenzyme (left) and RNase P holoenzyme-substrate complex (right) suggested by Tsai et al.[180] The helical regions of the RNA are numbered P1-P17.

The above-mentioned studies allow general placement of the protein subunit in structural models of RNase P, and identify two important regions of RNA-protein interaction involving: **i**) the highly conserved region (~11 amino acids ) in the α2 helix comprising basic amino acid residues, including the RNR motif, which contacts the P2, P3 and P4 helices of the catalytic domain (Figure 70); **ii**) a hydrophobic central cleft formed by the α1-helix and the face of the central β-sheet, which contacts the 5'-leader sequence of the ptRNA at positions at least 4nt from the cleavage site, but does not interact directly with the RNase P RNA (Figure 71). Additionally, the so-called metal binding loop (between β2 and β3) containing a cluster of negatively charged residues is

proposed to be a third RNA-binding region. The acidic residues in this exposed region could mediate RNA contacts via co-ordinated metal ions.[168] Deletion studies revealed that the C-terminal region and the RNR motif of the C5 protein are involved in binding to the RNA subunit of RNase P from *E. coli* (commonly named as M1 RNA). On the other hand, N-terminal deletions retained the M1 RNA binding ability. In spite of the lack of M1 RNA binding ability in the C-terminal deletion derivatives, they promote RNase P catalysis just like the N-terminal deletion derivatives.[182] Very recently Niranjanakumari[181] suggested that the RNR motif of the RNase P protein is located near (<15Å) the ptRNA cleavage site, the base of the ptRNA acceptor stem and the helix P4 of RNase P RNA. The P4 helix is a highly conserved RNA structure thought to be a part of the active site in the RNase P-ptRNA complex.

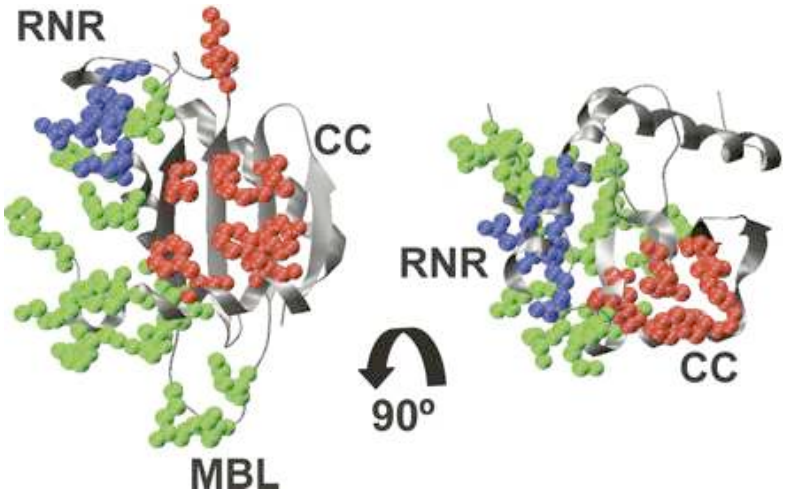



**Figure 71.** Three-dimensional structure of *B. subtilis* RNase P. Summary of cross-linking and AOP (acetomido-1,10-*o*-phenathroline) cleavage experiments presented by Niranjanakumari et al.[181] Labels indicate locations of the RNR motif, the metal binding loop (MBL) and the central cleft (CC). Highlighted amino acids indicate positions where cross-linking/cleavage occur to RNase P RNA only (green), to ptRNA only (red), or to both (blue).

Functional roles assigned to the basic bacterial RNase P protein subunit in general include: **i)** electrostatic neutralisation effects on the RNA domain and 5'-ptRNA backbone phosphates, **ii)** conformational stabilisation and fine-tuning of the active site structure, **iii)** increase in the binding affinity for ptRNA, which is linked to an increase in affinity for key metal ions involved in ptRNA binding and catalysis, and **iv)** mediation of RNase P dimer formation.[183]

In agreement with previously mentioned biochemical studies showing that the *B. subtilis* protein can substitute for the protein component of *E. coli* RNase P, both *in vivo* and *in vitro*, the cross-linking and cleavage studies showed similar patterns of

interactions indicating that the position and orientation of the P protein is similar in both holoenzymes.[156,158,181]

### 4.1.2. RNase P: A promising molecular target for the development of new drugs

The recent increase in drug-resistant bacteria and pathogenic bacterial infections has motivated intense efforts to develop novel antibacterial compounds. For this, the selection of effective molecular targets is of utmost importance. Bacterial RNase P combines several favourable target features. It contains an essential catalytic RNA subunit. The RNase P is a highly efficient enzyme, with a generally low cellular abundance. Furthermore, the differences in structure between bacterial and eucaryotic RNase P should allow effective discrimination between humans and the bacterial pathogens.

Some examples of RNase P inhibitors have already been discovered. Aminoglycosides, such as neomycin B, kanamycin, gentamycin have been reported to inhibit RNase P activity by a non-competitive mode of action.[184,185] The studies on aminoglysides suggest a competition between $Mg^{2+}$ ions and aminoglycoside antibiotic for the same binding sites on the enzyme. Gopalan and co-workers prepared aminoglycoside-arginine conjugates (NeoR), which combined the two different functionalities capable of RNA binding, and display a 500-fold increase in inhibitory efficiency ($IC_{50}$=60 μM versus 125 nM for neomycin B and NeoR, respectively) and good discrimination between bacterial and human RNase P.[186] Another approach was taken by Hori et al.[187-189] by designing benzimidazole-derivatives and porphines/porphyrins, which act by binding to the substrate ptRNA. Bacterial RNase P seems to be also an attractive target for antisense-based strategies, as reported by Willkomm.[190]

### 4.1.3. Project outline

The three-dimensional models of RNase P holoenzymes of *B. subtilis* and *E. coli*, which are based on extensive cross-linking and cleavage studies discussed above, have focused attention on the possible design of inhibitors of the RNA-protein interaction in this ribonuclease. The structural information available up-to-date shows the importance of the region containing the RNR-motif in helix α2 of RNase P protein, since it most probably contacts the catalytic domain of RNase P RNA. It was envisioned here, that it might be possible to mimic the α-helix of this critical region using a β-hairpin peptidomimetic and so to design novel inhibitors of this protein-RNA interaction in RNase P (Figure 72).

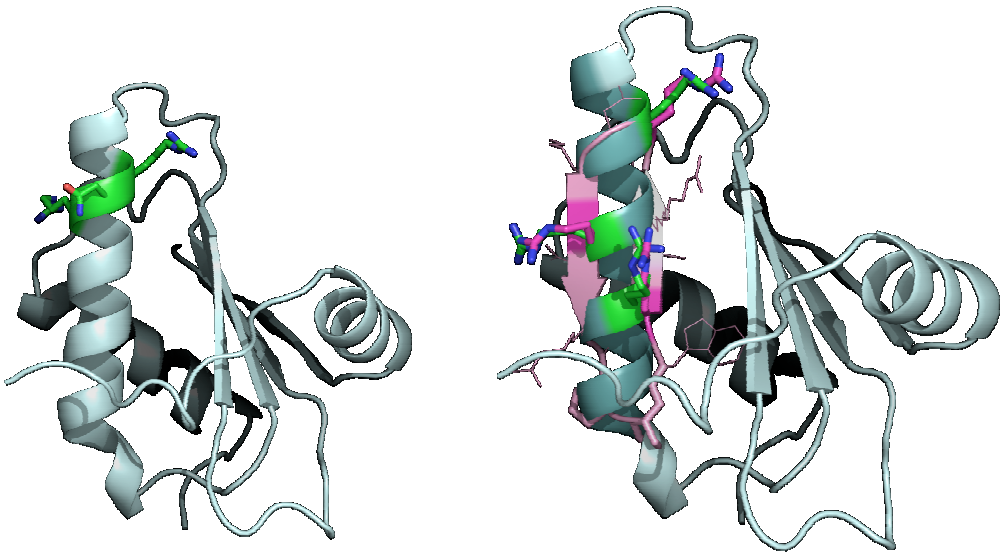


**Figure 72.** *Left:* Structure of the protein unit of RNase P form *B. subtilis* (**PDB 1A6F**)[168] with the α2-helix at the front and the RNR motif highlighted (green, side chains shown); *right:* the superposition of a β-hairpin peptide (pink) on the RNR motif. The side chains of Arg60, Arg65 and Arg68 are shown (from top to bottom).

It was reasoned that first assaying the existing library of Rev-RRE inhibitors, that share sequence similarities to the RNA-binding domain of RNase P protein subunit might deliver some preliminary leads. Furthermore, a library of mimetics might be designed to specifically target the RNase P. It was envisioned that this strategy might lead to the discovery of a novel class of RNase P inhibitors with potential utility in the search for new antibacterial agents.

## 4.2. Results

### 4.2.1. Design of novel RNase P protein-RNA interaction inhibitors

Encouraged by recent success in using a β-hairpin structure to mimic an α-helical RNA-binding domain of Rev protein and inhibit the Rev-RRE RNA interaction, we explored here a new approach to inhibitors of the protein-RNA interaction in RNase P, based on conformationally constrained β-hairpin peptidomimetics. From structural studies, it was known that the α2-helix of the protein subunit plays a critical role in contacting its cognate RNA. We assumed that a β-hairpin peptide might provide a robust scaffold upon which the groups critical for RNase P RNA recognition could be displayed. The α-helical structure of the basic domain of the protein, containing the highly conserved RNR motif, presents the energetically important residues Phe71-Arg60 to the RNA (Figure 73, left). According to this strategy, the hairpin mimetics should contain a twelve amino acid loop mounted on the D-Pro-L-Pro dipeptide template.

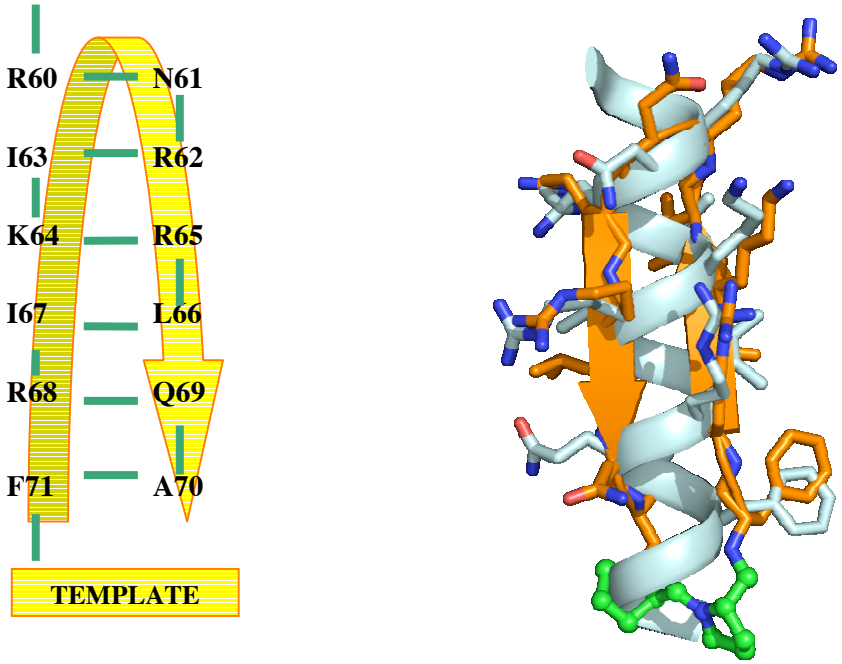



**Figure 73.** Schematic representation of the amino acids in the α2-helix of RNase P protein subunit from *B. subtilis* (Arg60-Phe71) overlaid with a β-hairpin peptidomimetic (Template+ yellow arrow). Model of the template-bound β-hairpin peptidomimetic (orange) designed to mimic the α-helical RNA-binding domain of RNase P protein (grey).

The computational model (Figure 73, right) of the β-hairpin superimposed on the known crystal structure of the α2-helix from *B. subtilis* RNase P protein, suggests a good overlap of the amino acids side chains. The hairpin could act as a scaffold to pre-organise these side chains to give a geometry similar to that seen in the helical motif.

To test this hypothesis, we first assayed a small family of cyclic β-hairpin peptidomimetics (**R-01 to R-27**), prepared in earlier work, which fulfil the structural requirements and display similarities with the target sequence, for their ability to mimic the α2-helix of RNase P protein and to bind to the target RNA.

### 4.2.2. Antimicrobial activity and selectivity of peptides R-01—R-27

The **R-01—R-27** peptidomimetics (from the R1 & R2 libraries) were tested first against *E. coli* to check their antimicrobial activity and then against other bacterial strains, including gram-positive bacteria, to investigate their selectivity (Table 36).

Among the 27 compounds assayed, two **R-17** and **R-22**, exhibit high antimicrobial activity towards *E. coli*, while being completely inactive against the other bacterial strains tested. The only difference between these peptides is at positions 7 and 8, where two amino acids Lys and Gly are swapped (Table 36). From an analysis of the inactive sequences, it appears that Trp1, Leu3, $^{D}$Arg6 are important for the activity of these peptides.

**Table 36.** The **R-01—R-27** library. Residues 1 and 12 are attached to the D-Pro-L-Pro template. MIC values (in μg/ml) were measured against the bacterial strains indicated.

| Mimetic | Position | | | | | | | | | | | | MIC[μg/ml] *E. coli* ATCC25922 | MIC[μg/ml] *P. aeruginosa* ATCC27853 | MIC[μg/ml] *P. aeruginosa* PAOI | MIC[μg/ml] *S. aureus* ATCC29213 | MIC[μg/ml] *S. aureus* ATCC25923 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *1* | 2 | *3* | 4 | *5* | 6 | *7* | 8 | *9* | 10 | *11* | 12 | | | | | |
| **R-01** | W | R | R | R | A | T | $^{D}$R | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-02** | W | R | R | R | A | P | $^{D}$R | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-03** | W | R | R | R | G | P | $^{D}$R | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-04** | W | R | R | R | V | P | $^{D}$R | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-05** | W | R | R | R | A | G | $^{D}$R | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-06** | W | R | R | R | A | G | K | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-07** | W | R | R | R | A | T | $^{D}$R | G | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-08** | W | R | R | R | A | P | $^{D}$R | G | R | N | R | R | 32 | >64 | >64 | >64 | >64 |
| **R-09** | W | R | R | R | G | P | $^{D}$R | G | R | N | R | R | 16 | >64 | >64 | >64 | >64 |
| **R-10** | W | R | R | R | G | T | $^{D}$R | Q | R | N | R | R | 32 | >64 | >64 | >64 | >64 |
| **R-11** | W | R | R | R | V | G | $^{D}$R | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-12** | W | R | R | R | A | S | $^{D}$R | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-13** | W | R | R | R | A | K | G | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-14** | W | R | L | R | A | K | G | Q | R | N | R | R | 32 | 64 | >64 | >64 | >64 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **R-15** | W | R | Q | R | A | K | G | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-16** | W | R | R | R | A | K | G | Q | R | N | R | V | >64 | >64 | >64 | >64 | >64 |
| **R-17** | **W** | **R** | **L** | **R** | **A** | **$^{D}$R** | **K** | **G** | **R** | **N** | **R** | **R** | **4** | **>64** | **>64** | **>64** | **>64** |
| **R-18** | W | R | L | R | A | K | K | G | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-19** | W | R | K | R | A | G | K | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-20** | W | R | K | R | A | K | G | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-21** | W | R | R | R | A | $^{D}$R | G | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-22** | **W** | **R** | **L** | **R** | **A** | **$^{D}$R** | **G** | **K** | **R** | **N** | **R** | **R** | **4** | **>64** | **>64** | **>64** | **>64** |
| **R-23** | R | R | L | R | A | $^{D}$R | G | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-24** | R | R | L | R | A | K | G | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-25** | R | V | R | R | A | K | G | Q | R | N | R | R | >64 | >64 | >64 | >64 | >64 |
| **R-26** | R | C | R | R | A | K | G | Q | R | R | C | R | >64 | >64 | >64 | >64 | >64 |
| **R-27** | K | R | Q | R | T | K | G | R | R | L | O | R | >64 | >64 | >64 | >64 | >64 |

### 4.2.3. Biological activities and selectivities

The primary criterion for the evaluation of antimicrobial peptides (AMPs) is their activity against micro-organisms. The RNase P peptidomimetics were assayed for antimicrobial activity using the NCCLS broth microdilution method. Briefly, two-fold serial dilutions of AMPs were prepared in Mueller-Hinton (MH) broth and inoculated with approximately $5x10^5$ colony forming units (CFU/ml) of the appropriate organism. The **minimal inhibitory concentrations (MIC)** are reported as those concentrations of test compound that completely inhibit bacterial growth after an incubation period of 20 h, as judged by visual inspection. The MIC of each mimetic was measured against the target organism *Escherichia coli* ATCC25922. Some peptidomimetics were assayed with other strains of *Escherichia coli*, as well as other micro-organisms, including *Pseudomonas aeruginosa* and *Pseudomonas putida*, representing Gram-negative bacteria, and *Staphylococcus aureus* and *Bacillus subtilis*, representing Gram-positive bacteria.

### 4.2.4. *In vitro* assay for inhibition of *Escherichia coli* RNase P

The *in vitro* assay for inhibition of *Escherichia coli* RNase P was performed in cooperation with Professor V. Gopalan at the Department of Biochemistry at The Ohio State University. In preliminary studies with **R-01—R-27**, two peptides able to inhibit exclusively the growth of *E. coli* were discovered. In the next step, assays were performed to determine if any of these mimetics are able to inhibit RNase P *in vitro*.

Precursor tRNA was incubated with RNase P of *E. coli* and with the tested peptide. As a negative control only ptRNA was present in the assay, and as positive control the assay mixture contained no peptide inhibitor. Upon processing, the radiolabelled ptRNA is cleaved into mature tRNA and 5'-leader, both of which also contain radioactivity. The electrophoretic properties of the precursor, mature tRNA and 5'-leader differ, so by EMSA it is possible to follow the course of the processing by RNase P (Figure 74).

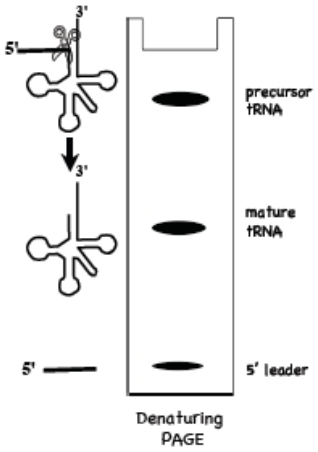



**Figure 74.** Principle of the RNase P-mediated processing of precursor tRNA, as monitored by EMSA.

The initial screen was performed with the selected peptides at a concentration of 5μM (Figure 75). The results of this initial EMSA are in close agreement with the information obtained from the antimicrobial assay. Peptides **R-17** and **R-22** show an inhibitory effect on RNase P from *E. coli*. The processing of the ptRNA takes place to a smaller extent in the presence of these two peptides, in comparison to the control lane and those for the other peptides.

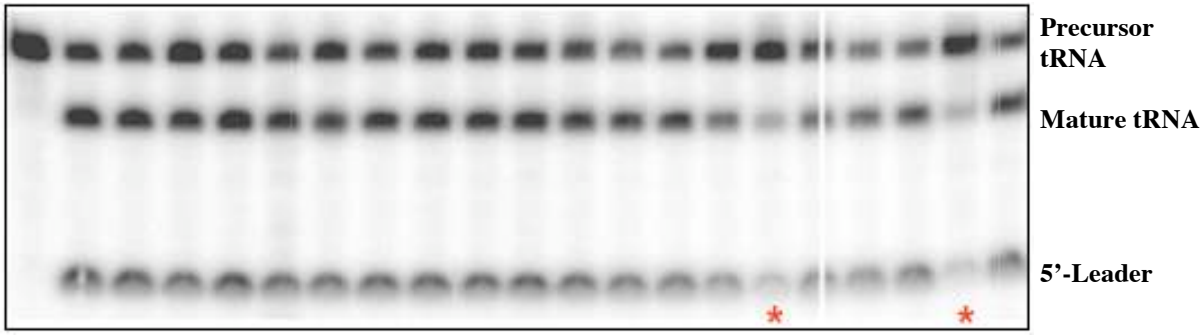



**Figure 75.** Initial EMSA assay of the inhibitory potential of the selected peptides, performed with *E. coli* RNase P and precursor tRNA. Prod = control without inhibitor present.

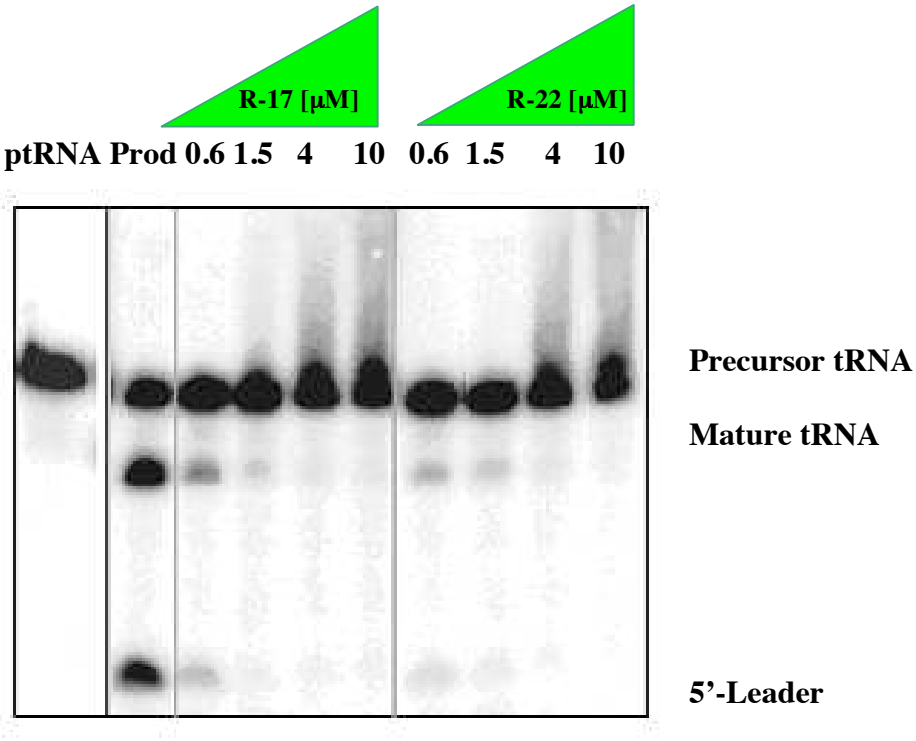



**Figure 76.** EMSA with a series of concentrations of **R-17** and **R-22** peptides. Assays were performed with RNase P from *E. coli*. Concentrations of cations: 200 mM $NH_4^+$ and 5 mM $Mg^{2+}$. Prod = control without inhibitor present.

The concentration-dependence of the inhibition by these two lead peptides was also investigated (Figure 76). Under the assay conditions, both peptides inhibit completely maturation of precursor tRNA at a concentration of 4 μM. The MIC value for these peptides is ≈ 4 μg/ml or approximately 2.2 μM. These values are very similar, which is at least consistent with the antimicrobial activity arising due to inhibition of RNase P.

### 4.2.5. Alanine scan library P1

Mimetics **R-17** and **R-22** are able to inhibit growth of *E. coli* with MIC values of ≈ 4 μg/ml. In order to identify the residues that are energetically important for antimicrobial activity of the peptides, a library of 11 mimetics was prepared, in which each residue in the sequence (R-22 was chosen) was replaced in turn by Ala (Table 37).

The RNase P mimetics were synthesized, using methods described previously in Paragraph 2.2.2. In a typical synthesis, the main component in the crude product was the desired RNase P mimetic. After purification by HPLC, the desired peptidomimetics

were typically obtained in 20-40% overall yield, and each was >95% pure as determined by analytical HPLC and gave ESI MS data consistent with the calculated masses (see Appendix 4 for analytical data).

The MIC values for the Ala-scanning library (Table 37) revealed that the replacement of Trp1 and Leu3 results in complete loss in activity in the range of concentration used. The $^{D}$Arg6Ala mutation produced a compound of 8-fold lower activity than the lead mimetic. Only mutations at positions 7 (Gly) and 10 (Asn) are unaffected in terms of activity. These findings support our design strategy, as in the model position 1 is occupied by Phe (Trp in **R-22**), at position 3 Ile is present (Leu in **R-22**), and position 6 is Arg (D-Arg in **R-22**).

It seems that positions 1 and 3 have to be occupied by aromatic/large aliphatic groups, while Arg is important for the turn region. Reducing the positive charge of the mimetic most of the time has only a small influence on the activity (mimetics **P1-08**, **P1-11**, **P1-12**). Surprisingly, replacement of the polar residue Asn10 by Ala does not bring any changes in activity. This indicates that position 10 may accommodate residues of having different side chain properties.

**Table 37**. Alanine scanning library. Residues 1 and 12 are attached to a D-Pro-L-Pro template. MIC values (in μg/ml) of AMPs were measured for 6 strains of bacteria.

| **Mimetic** | **Position** | | | | | | | | | | | | ***E. coli*** **ATCC 25922** | ***B. subtilis*** **3256** | ***P. aeruginosa*** **ATCC 27853** | ***P. aeruginosa*** **PAOI** | ***S. aureus*** **ATCC 29213** | ***S. aureus*** **ATCC 25923** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | | | | | | |
| **R-22** | **W** | **R** | **L** | **R** | **A** | **$^{D}$R** | **G** | **K** | **R** | **N** | **R** | **R** | **4** | **3** | **>64** | **>64** | **>64** | **>64** |
| **P1-01** | **A** | R | L | R | A | $^{D}$R | G | K | R | N | R | R | **>64** | 64 | >64 | >64 | >64 | >64 |
| **P1-02** | W | **A** | L | R | A | $^{D}$R | G | K | R | N | R | R | 16 | 16 | >64 | >64 | >64 | >64 |
| **P1-03** | W | R | **A** | R | A | $^{D}$R | G | K | R | N | R | R | **>64** | 24 | >64 | >64 | >64 | >64 |
| **P1-04** | W | R | L | **A** | A | $^{D}$R | G | K | R | N | R | R | 16 | 24 | >64 | >64 | >64 | >64 |
| **P1-05** | W | R | L | R | A | **A** | G | K | R | N | R | R | **32** | 24 | >64 | >64 | >64 | >64 |
| **P1-06** | W | R | L | R | A | $^{D}$R | **A** | K | R | N | R | R | **4** | 16 | >64 | >64 | >64 | >64 |
| **P1-07** | W | R | L | R | A | $^{D}$R | G | **A** | R | N | R | R | **8** | 16 | >64 | >64 | >64 | >64 |
| **P1-08** | W | R | L | R | A | $^{D}$R | G | K | **A** | N | R | R | **8** | 32 | >64 | >64 | >64 | >64 |
| **P1-09** | W | R | L | R | A | $^{D}$R | G | K | R | **A** | R | R | **4** | 6 | >64 | >64 | >64 | >64 |
| **P1-10** | W | R | L | R | A | $^{D}$R | G | K | R | N | **A** | R | **16** | 64 | >64 | >64 | >64 | >64 |
| **P1-11** | W | R | L | R | A | $^{D}$R | G | K | R | N | R | **A** | **8** | 12 | >64 | >64 | >64 | >64 |

### 4.2.6. Mimetic library P2: Attempts to optimise the lead compound

Following the observations from the initial screen (peptides R-01-R-27) and Ala-scanning of R-22, a new library was designed, in order to introduce amino acid residues that are present in the β-hairpin model (Figure 73). Previous results highlighted the importance of the Trp1, Leu3 and D-Arg6, therefore in this library mutations other than Ala were introduced at these positions. It was also established that the charge of peptides can be reduced, at least from +7 to +6. In the present approach some peptides have a total charge even lower (e.g. +4 in P2-22 and +3 in P2-24). P2-11 represents a sequence based entirely on the theoretical model (Figure 73). In other peptides single/multiple changes are introduced to follow the model sequence (P2-03—P2-10). Furthermore, all Arg residues were mutated in turn into Lys in order to determine the importance of the guanidine group (P2-12—P2-17). In one of the peptides, P2-21, a disulfide bridge was introduced at non-hydrogen bonding positions 2 and 11.

These peptides were prepared using methods described previously. After purification by HPLC, the desired peptidomimetics were typically obtained in 20-40% overall yield, and each was >95% pure as determined by analytical HPLC and gave ESI-MS data consistent with the calculated masses (see Appendix 4 for analytical data).

The antimicrobial activity of these peptides is summarised in the Table 38. This optimisation cycle produced many mimetics having improved activity (MIC <1 μg/ml) and very high selectivity for *E. coli*. The most potent peptidomimetic, **P2-05,** exhibits an MIC of 0.06 μg/ml. This new lead compound represents a sequence that is a chimera between **R-22** and the model sequence presented in **P2-11**. It appears, that the mutations Ala5Ile and Asn10Leu are of major importance for the antimicrobial activity. The results show that position 1 can accommodate neutral aromatic (Phe or Trp) or large aliphatic residues (Ile), however, Trp1Tyr brings a 16-fold loss of activity (**P2-04** vs. **P2-10**). This observation suggests that at this position large neutral amino acid residues, preferentially aromatic, are optimal, while introducing polarity possibly disturbs hydrophobic interactions.

The results concerning mutations at position 6 in the turn region show that the stereochemistry at this position is more important than the character of the amino acid. The mutation $^{D}$Arg6Arg brings a 32-fold decrease in activity (**P2-04** vs. **P2-08**), while the mutation $^{D}$Arg6$^{D}$Lys produces a peptide with retained activity (**P2-04** vs. **P2-14**).

Peptides **P2-12—P2-17** show that the replacement of Arg by Lys does not influence the activity of the peptides, therefore one can assume that the guanidinium groups are not critical for the interaction with the target. The mutations of Arg/Lys to Gln, which brings a reduction of the charge, when introduced singly, are very well tolerated (**P2-06**, **P2-19**, **P2-20**), but in multiple copies (**P2-22**, **P2-23**, **P2-24**), they bring a loss in activity. It is noteworthy, that reducing the charge from +7 to +5 (peptides **P2-11** and **P2-07**) does not have any influence on the peptide potency. This observation may be very important for further optimisation in terms of improving other properties and plasma stability.

It is noteworthy, that all the peptides active against *E. coli* show also some, but strongly reduced, antimicrobial activity against *B. subtilis*. Interestingly, **P2-11** representing the model sequence based on the RNA-binding domain of the *B. subtilis* protein subunit, is the only peptidomimetic that inhibits more strongly growth of *B. subtilis* than of *E. coli* (0.12 vs 0.37 μg/ml). In addition, the peptides **P2-05** and **P2-11** exhibit some inhibitory potency against other bacterial strains tested, however, the MIC values are at least 10-fold higher.

**Table 38**. Library P2 of RNase P peptidomimetics. Residues 1 and 12 are attached to a D-Pro-L-Pro template. MIC values (in μg/ml) of AMPs were measured for 6 strains of bacteria.

| **Mimetic** | **Position** | | | | | | | | | | | | ***E. coli* ATCC 25922** | ***B. subtilis* 3256** | ***P. aeruginosa* ATCC 27853** | ***P. aeruginosa* PAOI** | **]*S. aureus* ATCC 29213** | ***S. aureus* ATCC 25923** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | | | | | | |
| **R-22** | **W** | **R** | **L** | **R** | **A** | **$^{D}$R** | **G** | **K** | **R** | **N** | **R** | **R** | **4** | **3** | **>64** | **>64** | **>64** | **>64** |
| **P2-01** | W | R | **I** | R | I | $^{D}$R | G | K | R | N | R | R | 0.5 | 4 | >64 | >64 | >64 | >64 |
| **P2-02** | W | R | **I** | R | A | $^{D}$R | G | K | R | N | R | R | 4 | 6 | >64 | >64 | >64 | >64 |
| **P2-03** | W | R | L | R | **I** | $^{D}$R | G | K | R | N | R | R | 0.5 | 2 | >64 | >64 | >64 | >64 |
| **P2-04** | W | R | **I** | R | **I** | $^{D}$R | **N** | K | R | N | R | R | 0.5 | 2 | >64 | >64 | >64 | >64 |
| **P2-05** | W | R | **I** | R | **I** | $^{D}$R | **N** | K | R | **L** | R | R | **0.06** | 0.25 | 16 | 32 | 4 | 8 |
| **P2-06** | W | R | **I** | R | **I** | $^{D}$R | **N** | K | R | N | **Q** | R | 0.5 | 6 | >64 | >64 | >64 | >64 |
| **P2-07** | W | R | **I** | R | **I** | $^{D}$R | **N** | K | R | N | **Q** | **A** | 0.5 | 1.5 | 64 | >64 | >64 | >64 |
| **P2-08** | W | R | **I** | R | **I** | **R** | **N** | K | R | N | R | R | 16 | 6 | >64 | >64 | >64 | >64 |
| **P2-09** | **F** | R | **I** | R | **I** | $^{D}$R | **N** | K | R | N | R | R | 0.25 | 0.75 | >64 | 64 | 32 | 32 |
| **P2-10** | **Y** | R | **I** | R | **I** | $^{D}$R | **N** | K | R | N | R | R | 8 | 12 | >64 | >64 | >64 | >64 |
| **P2-11** | **F** | R | **I** | **K** | **I** | $^{D}$R | **N** | K | R | **L** | **Q** | **A** | 0.37 | 0.12 | 4 | 4 | 4 | 8 |
| **P2-12** | W | **K** | **I** | R | **I** | $^{D}$R | **N** | K | R | N | R | R | 0.75 | 1.5 | >64 | >64 | >64 | >64 |
| **P2-13** | W | R | **I** | **K** | **I** | $^{D}$R | **N** | K | R | N | R | R | 0.75 | 3 | >64 | >64 | >64 | >64 |
| **P2-14** | W | R | **I** | R | **I** | **$^{D}$K** | **N** | K | R | N | R | R | 0.5 | 6 | >64 | >64 | >64 | >64 |
| **P2-15** | W | R | **I** | R | **I** | $^{D}$R | **N** | K | **K** | N | R | R | 0.12 | 2 | >64 | >64 | >64 | >64 |
| **P2-16** | W | R | **I** | R | **I** | $^{D}$R | **N** | K | R | N | **K** | R | 0.5 | 4 | >64 | >64 | >64 | >64 |
| **P2-17** | W | R | **I** | R | **I** | $^{D}$R | **N** | K | R | N | R | **K** | 0.25 | 6 | >64 | >64 | >64 | >64 |
| **P2-18** | **I** | R | **I** | R | **I** | $^{D}$R | **N** | K | R | N | R | R | 0.75 | 4 | >64 | >64 | >64 | >64 |
| **P2-19** | W | **Q** | **I** | R | **I** | $^{D}$R | **N** | K | R | N | R | R | 0.37 | 8 | >64 | >64 | >64 | >64 |
| **P2-20** | W | R | **I** | R | **I** | $^{D}$R | **N** | **Q** | R | N | R | R | 0.25 | 8 | >64 | >64 | >64 | >64 |
| **P2-21** | W | **C** | **I** | R | **I** | $^{D}$R | **N** | K | R | N | **C** | R | 1.5 | 8 | >64 | >64 | >64 | >64 |
| **P2-22** | W | R | **I** | R | **I** | $^{D}$R | **N** | **Q** | R | N | **Q** | **A** | 4 | 16 | >64 | >64 | >64 | >64 |
| **P2-23** | W | R | **I** | R | **I** | $^{D}$R | **N** | **Q** | R | N | **Q** | R | 2 | 32 | >64 | >64 | >64 | >64 |
| **P2-24** | W | R | **I** | R | **I** | $^{D}$R | **N** | **Q** | **Q** | N | **Q** | **A** | >64 | >64 | >64 | >64 | >64 | >64 |

### 4.2.7. An evaluation of the importance of the D-Pro-L-Pro template

In order to evaluate the importance of the template, the linear version of one of the active peptidomimetics, namely **P2-04,** was prepared (Table 39).

**Table 39.** Sequences and MIC values for **P2-04** and its linear derivative. Note, the N-terminus of **P2-04a** is acetylated, and the C-terminus is an amide.

| Mimetic | Position | | | | | | | | | | | | Template | MIC[μg/ml] *E. coli* ATCC25922 | MIC[μg/ml] *B. subtilis* 3256 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *1* | 2 | *3* | 4 | *5* | 6 | *7* | 8 | *9* | 10 | *11* | 12 | | | |
| **P2-04** | W | R | I | R | I | $^{D}$R | N | K | R | N | R | R | $^{D}$P-P | 0.5 | 2 |
| **P2-04a** | Ac-W | R | I | R | I | $^{D}$R | N | K | R | N | R | R-$NH_2$ | -- | 16 | 4 |

The 32-fold loss in antimicrobial activity of the linear peptide against *E. coli* in comparison to the cyclic one, suggests that either the constrained β-hairpin conformation or the template itself (or both) might be important for the activity. In sharp contrast, the inhibitory potency of the linear **P2-04a** determined against *B. subtilis*, a closely related bacterial strain, is only slightly influenced by removal of the template.

### 4.2.8. The enantiomeric form of the peptide

In the past, enantiomeric forms of antimicrobial peptides with all-D-amino acids were widely used to study the mechanism(s) of antimicrobial action.[191,192] It has been suggested that the interactions between antimicrobial peptides and the cell membrane might not require from these class of peptides a specific chirality. Indeed, many studies have shown that for many membranolytic peptides the all-D-amino acid forms have equal activities to their all-L-enantiomers,[193,194] suggesting that the antimicrobial mechanism of these peptides does not involve a stereoselective interaction with a chiral enzyme or protein receptor. In addition, all-D-peptides are resistant to proteolytic enzyme degradation, which enhances their potential as clinical therapeutics. Similar conclusions have been drawn with ligand-RNA interaction inhibitors, where the D-forms of potent ligands exhibit only a slight decrease in affinities to the target RNAs, as discussed Paragraph 1.1.3.

However, there are some examples known of antimicrobial peptides from natural sources where the D-form is much less active than the L-form. These peptides typically do not show membranolytic activity, and must have a different mechanism of action.[195,196]

**Table 40.** Sequences and MIC values for **P2-04** and its enantiomeric form.

| Mimetic | Position | | | | | | | | | | | | TEMPLATE | MIC[μg/ml] *E. coli* ATCC25922 | MIC[μg/ml] *B. subtilis* 3256 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *1* | 2 | *3* | 4 | *5* | 6 | *7* | 8 | *9* | 10 | *11* | 12 | | | |
| **P2-04** | W | R | I | R | I | $^{D}$R | N | K | R | N | R | R | $^{D}$P-P | 0.5 | 2 |
| **P2-04b** | $^{D}$W | $^{D}$R | $^{D}$I | $^{D}$R | $^{D}$I | R | $^{D}$N | $^{D}$K | $^{D}$R | $^{D}$L | $^{D}$R | $^{D}$R | P-$^{D}$P | 0.5 | 0.2 |

The results of the antimicrobial assay (Table 40) show that the enantiomeric form of **P2-04** has the same activity as the parent peptide against *E. coli*. Once again there is a clear difference in behaviour between *E. coli* and *B. subtilis*. **P2-04b** exhibits more than 10-fold higher inhibitory activity against *B. subtilis* than the parent peptide. It is also ≈3-fold more active against *B. subtilis* than *E. coli*.

### 4.2.9. Mimetic library P3: Further optimisation of the lead compound

The results form library **P2** have shown that one strand of the β-hairpin (amino acids 1-6) is much more sensitive to changes than the other. In this library, the beneficial changes were combined in order to enhance the activity, as well as to decrease the total charge of the peptide.

So far the model of the β-hairpin peptide was based on the sequence derived from *B. subtilis*, since the crystal structure of the RNA and the protein subunits are available. In this library, the sequence derived from the *E. coli* RNA-binding domain has been tested in order to evaluate if this approach would bring an improvement in activity and specificity against this strain of bacteria. The peptide derived from the sequence of the *E. coli* RNase P subunit, using the model shown in Figure 73, is represented by **P3-04** (Table 41).

**Table 41.** Library P3 of RNase P peptidomimetics. MIC values (in μg/ml) of AMPs were measured against *E. coli* ATCC25922.

| **Mimetic** | **Position** | | | | | | | | | | | | **MIC [μg/ml] *E. coli* ATCC25922** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | |
| **R-22** | **W** | **R** | **L** | **R** | **A** | **$^{D}$R** | **G** | **K** | **R** | **N** | **R** | **R** | **4** |
| **P2-05** | **W** | **R** | **I** | **R** | **I** | **$^{D}$R** | **N** | **K** | **R** | **L** | **R** | **R** | **0.06** |
| **P3-01** | W | R | I | R | I | $^{D}$R | N | **Q** | **K** | L | R | R | 0.37 |
| **P3-02** | W | R | I | R | I | $^{D}$R | N | **Q** | **K** | **I** | R | R | 0.12 |
| **P3-03** | W | R | I | R | I | $^{D}$R | N | K | **O** | L | R | R | 0.12 |
| **P3-04** | **F** | R | **T** | R | I | $^{D}$R | N | K | R | L | **E** | **S** | 16 |
| **P3-05** | **F** | R | I | R | I | $^{D}$R | N | K | R | L | R | R | 0.09 |
| **P3-06** | W | R | I | R | I | $^{D}$R | N | K | R | L | R | **I** | 2 |
| **P3-07** | **F** | R | I | R | I | $^{D}$R | N | K | R | L | R | **S** | 0.5 |
| **P3-08** | W | R | I | R | I | $^{D}$R | N | K | R | L | **Q** | **A** | 0.25 |
| **P3-09** | **Cha** | R | I | R | I | $^{D}$R | N | K | R | L | R | R | 0.12 |
| **P3-10** | W | R | I | R | I | $^{D}$R | N | **Q** | R | L | **Q** | **A** | 0.75 |

The results of MIC assays for the library **P3** are reported in Table 41. Among the peptides tested no further gain in activity was achieved, however, almost all of the peptides exhibit very high potency. **P2-05** represents so far the optimal sequence for inhibitory activity against *E. coli*. In agreement with previous studies, position 1 appears to tolerate large aliphatic residues (as seen in **P3-09,** Trp1Cha). The incorporation of Leu at position 10 allows reduction of the total charge to +4 with much lower loss in activity compared to that seen in the previous library (i.e. 0.75 μg/ml vs. 4 μg/ml in **P3-10** and **P2-22**, respectively). Reduction of the total charge of the peptides might be important when considering the pharmacological properties of the compounds. Highly positively charged molecules are often found to be toxic (due to haemolytic activity and histamine release associated with their cationic nature). In addition, a replacement of the basic amino acid residues might also bring improvements in the stability against proteolytic degradation.

It is noteworthy, that **P3-04**, whose sequence is derived from the RNA-binding domain of *E. coli* RNase P exhibits a dramatic loss in antimicrobial activity in comparison to the most active peptidomimetic **P2-05** so-far (16 vs. 0.06 μg/ml). This suggests that the model of the RNA-binding domain derived from *B. subtilis* has more predictive value than the one for *E. coli*, perhaps due to the more detailed structural information available for *B. subtilis*. However, it might be that incorporation of just one

mutation (e.g. Thr3 or Glu11) has significant impact on activity, therefore further SAR studies are necessary.

### 4.2.10. Effect of divalent cations ($Ca^{2+}$ and $Mg^{2+}$) on antimicrobial activity

Many cationic antimicrobial peptides exhibit reduced potency in the presence of high concentrations of salt and, in particular, divalent cations. The $Ca^{2+}$ concentration in the extracellular fluid is kept at approximately $10^{-3}$ M, and its concentration inside the cells is at approximately $10^{-6}$ M. The $Mg^{2+}$ concentrations extracellular and intracellular are approximately 0.7-1.2 mM and 0.5 mM, respectively. Calcium remains almost wholly in plasma and very little in the cells, while magnesium is mostly found in the cells and less in plasma. It is thought that the presence of cations in the buffer reduces the affinity of cationic peptides to the negatively charged microbial membrane.[197,198] In order to evaluate the effect of $Ca^{2+}$ and $Mg^{2+}$ on inhibitory activity, two peptides **P2-04** and **P2-22** were tested.

**Table 42.** Summary of MIC values [μg/ml] for the assays performed in the presence of different concentrations of divalent cations. The peptides were assayed with the concentrations of calcium chloride and magnesium sulfate indicated or with a mixture of them both.

| **Mimetic** | **P2-04** | **P2-22** |
|---|---|---|
| **Concentration $Ca^{2+}$ /$Mg^{2+}$** | **MIC [μg/ml]** | |
| **0mM $Ca^{2+}$ / $Mg^{2+}$** | **0.5** | **4** |
| 0.5mM $CaCl_2$ | 0.5 | 12 |
| 2mM $CaCl_2$ | 4 | >64 |
| 5mM $CaCl_2$ | 16 | >64 |
| 10mM $CaCl_2$ | 64 | >64 |
| 20mM $CaCl_2$ | >64 | >64 |
| 40mM $CaCl_2$ | >64 | >64 |
| 5mM $MgSO_4$ | 3 | >64 |
| 20mM Mg $SO_4$ | 64 | >64 |
| 40mM Mg $SO_4$ | >64 | >64 |
| 5mM $CaCl_2$/Mg $SO_4$ | 64 | >64 |
| 20mM $CaCl_2$/Mg $SO_4$ | >64 | >64 |
| 40mM $CaCl_2$/Mg $SO_4$ | >64 | >64 |

The results of the assays with defined concentrations of divalent cations are reported in Table 42. Both peptides appear to be very sensitive to the presence of $Ca^{2+}$

and $Mg^{2+}$. However, the effect is dependent on the total positive charge of the peptide; **P2-04** has +7 positively charged guanidinium/amine groups under physiological conditions and is affected to lesser extent by the presence of cations than **P2-22** with 4 positive charges.

This experiment shows that the RNase P mimetics share this sensitivity to divalent cations with other known cationic antimicrobial peptides, e.g. defensins.[198] The sensitivity to cations might be a limiting factor for the potential use of this class of mimetics for pharmaceutical purposes. They could be effective in killing pathogens only at sites with low concentration of divalent cations.

### 4.2.11. Activity of RNase P peptidomimetics on a selection of *E. coli* strains

The previous experiments show that the peptidomimetics exhibit high antimicrobial activity and selectivity against an *E. coli* strain used commonly in laboratories, *E. coli* ATCC25922. To prove the ability to inhibit bacterial growth of pathogenic strains, 6 clinical isolates of *E. coli* were also tested in assays (Table 43).

**Table 43.** Susceptibility of pathogenic strains of *E. coli* to selected RNase P peptidomimetics. MIC values are given in μg/ml.

| Mimetic | *E. coli* ATCC25922 | *E. coli* 2138E/2151 | *E. coli* 2139E/2152 | *E. coli* 2140E/2153 | *E. coli* 2143E/2154 | *E. coli* 2144E/2155 | *E. coli* 3459E/2150 |
|---|---|---|---|---|---|---|---|
| **R-17** | **4** | 8 | 16 | 24 | 16 | 3 | 32 |
| **R-22** | **4** | 4 | 16 | 16 | 16 | 3 | 32 |
| **P1-01** | **>64** | >64 | >64 | >64 | >64 | >64 | >64 |
| **P1-06** | **4** | 4 | 6 | 8 | 8 | 6 | 8 |
| **P2-01** | **0.5** | 0.25 | 0.5 | 0.75 | 0.25 | 3 | 0.375 |
| **P2-03** | **0.5** | 0.375 | 0.25 | 0.375 | 0.125 | 0.75 | 0.25 |
| **P2-04** | **0.125** | 0.125 | 0.125 | 0.125 | 0.09 | 0.75 | 0.125 |
| **P3-10** | **0.75** | 1 | 1 | 1.5 | 0.5 | 0.5 | 1 |

These results prove that the peptidomimetics tested are able to inhibit bacterial growth of several pathogenic *E. coli* strains present in the clinic.

### 4.2.12. Haemolytic activity of RNase P peptidomimetics

In addition to antimicrobial activity, measured in MIC values, a second important criteria for the AMPs, with respect to their potential usefulness as therapeutic agents, is their haemolytic activity, i.e. the ability to lyse red blood cells. Such activity is often used as an indicator for the toxicity of AMPs towards (mammalian) host cells. Therefore, a low level of haemolysis is a pre-requisite for any systemic applications.

**Table 44.** Haemolytic activity (% hemolysis) of RNase P peptidomimetics and a control antimicrobial agent, protegrin-1 (PG-1), against human red blood cells, each at a concentration of 100 $\mu$g/ml.

| Mimetic | Haemolysis [%] |
|---|---|
| R-17 | 1.35 |
| R-22 | 0.5 |
| P1-01 | 0.07 |
| P1-02 | 0.57 |
| P1-03 | 0.1 |
| P1-04 | 0.5 |
| P1-05 | 0.5 |
| P1-06 | 0.45 |
| P1-07 | 0.4 |
| P1-08 | 0.45 |
| P1-09 | 0.83 |
| P1-10 | 0.5 |
| P1-11 | 0.9 |
| P2-01 | 1.27 |
| P2-02 | 0.85 |
| P2-03 | 1.37 |
| P2-04 | 1.37 |
| P2-05 | 1.9 |
| P2-06 | 0.95 |
| P2-07 | 1.83 |
| P2-08 | 1.0 |
| P2-09 | 1.47 |
| P2-10 | 0.65 |
| P2-11 | 3.15 |
| P2-12 | 0.93 |
| P2-13 | 1.05 |
| P2-14 | 1.15 |
| P2-15 | 1.05 |
| P2-16 | 1.10 |
| P2-17 | 0.7 |
| P2-18 | 0.9 |
| P2-19 | 1.4 |
| P2-20 | 1.0 |
| P2-21 | 0.7 |
| P2-22 | 1.93 |
| P2-23 | 1.3 |

| | |
|---|---|
| P2-24 | 1.5 |
| P2-04a | 1.13 |
| P2-04b | 1.3 |
| PG-1 | 45.35 |

The haemolytic activity of the RNase peptidomimetics was determined by their ability to lyse human red blood cells (hRBC), typically at peptide concentration of 100 μg/ml. The results of these experiments are expressed as percentage of haemolysis relative to 0.1% Triton-X 100$^{TM}$, which is taken as a reference corresponding to complete lysis of the blood cells. Protegrin-1 (PG-1), a natural antimicrobial peptide with known haemolytic properties was used as a reference. The results summarised in Table 44 show that the peptides have a very low level of haemolytic activity, which supports their usefulness for therapeutic purposes.

### 4.2.13. Plasma stability assays

The proteolytic stability of the compounds in biological settings is a major challenge for using peptides as prospective therapeutics in drug discovery programmes, as discussed previously in Paragraph 2.2.7. The peptides were assayed in human and mouse plasma as described in Paragraph 2.2.8.2. The results of the experiments are summarised in Table 45.

**Table 45.** Summary of the results from plasma stability assays. The half-life ($\tau_{1/2}$ [min]) and stability after 4 h (% remaining after 4h) are presented for each peptide after incubation with human or mouse plasma.

| **Mimetic** | **$\tau_{1/2}$ [min] in Human Plasma** | **Stability after 4 h in Human Plasma [%]** | **$\tau_{1/2}$ [min] in Mouse Plasma** | **Stability after 4h in Mouse Plasma [%]** |
|---|---|---|---|---|
| R-22 | 140 | 29.2 | 114 | 16.2 |
| P2-04 | >240 | 93.6 | 122 | 28.7 |
| P2-04b | >240 | 100 | >240 | 100 |
| P2-05 | >240 | 90.3 | >240 | 72.5 |
| P2-11 | >240 | 100 | >240 | 89.9 |

The results indicate that under the assay conditions all the peptides exhibit a high stability against proteolytic enzymes present in plasma. The peptides are more susceptible to degradation in mouse plasma than in human plasma, and similar results were obtained for the Tat peptidomimetics (Paragraph 2.2.8.2). As expected, the mostly-D-amino acid-containing **P2-04b** is completely stable in these experiments.

## 4.3. Discussion

The extensive structural, functional and sequence data available for the RNA and protein subunits of several bacterial RNase P ribonucleases, enabled in this work the design of a novel class of RNase P inhibitors. An initial assay of the β-hairpin peptides available from the Rev-RRE project delivered two compounds (**R-17** and **R-22**) with antimicrobial activity selective to *E. coli*. A similar observation was also made by EMSA with ptRNA and RNase P of *E. coli*. At the same time, a model β-hairpin peptide was designed and prepared by transpositioning the RNA-binding region displayed on an α-helical segment of the RNase P protein, onto a ß-hairpin scaffold to give mimetic **P2-11**.

The alanine scan of lead peptide **R-22**, allowed the identification of residues important for the inhibitory activity. Based on this information and on the sequence of **P2-11**, the step-wise optimisation of the lead peptidomimetic was carried out, through the design, screening, and SAR analysis of peptides libraries. The analysis of the SAR data highlighted the importance of Trp/Phe1, a large aliphatic group at position 3 (Leu or Ile) and the D-Arg6 in the turn region.

The most potent mimetic **P2-05** represents a chimera between sequences derived from the model peptide **P2-11** and **R-22** (Figure 77). Although **P2-11** represents an optimal model sequence, it is not as potent as the best mimetic. This indicates that although the approach to the design of the β-hairpin mimetics was successful, as long as a crystal or NMR structure of the *E. coli* RNase P RNA-protein complex is not available, it is difficult to rationally design improved mimetics of the RNA-binding domain.

Interestingly, only the model sequence derived using the *B. subtilis* RNase P, gave a peptide more active against *B. subtilis* than against *E. coli* (**P2-11**). Any other variants tested gave peptidomimetics with higher antimicrobial activity against *E. coli* than against *B. subtilis*. Using the sequence derived from the α2-helix of the protein domain of *E. coli* RNase P, **P3-04**, brought a 43-fold loss in activity in comparison to the model sequence (**P2-11**) from *B. subtilis*. This is again an indication that without detailed structural information on the complex from *E. coli*, it is difficult to rationally design inhibitors of this RNA-protein interaction. However, good progress has been made by exploiting this combinatorial approach to mimetic design and synthesis.

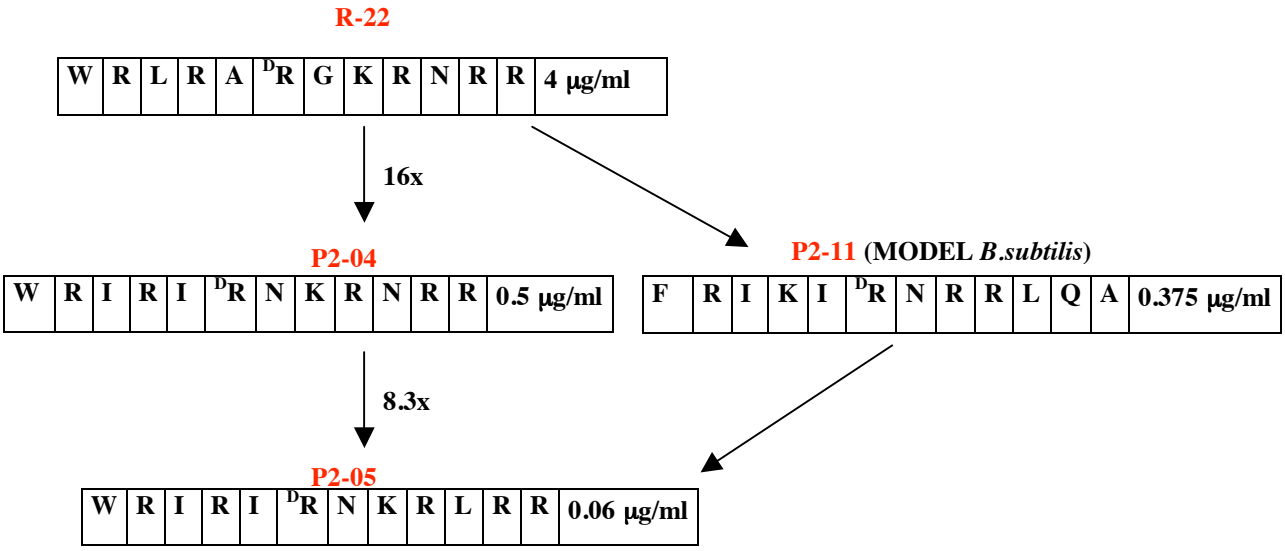



**Figure 77.** Representative examples of RNase P peptidomimetics, showing the stepwise improvements in antimicrobial activity against *E. coli*.

Decreasing the total positive charge of the peptides, performed in the **P2** and **P3** libraries, leads to the conclusion that in the case of these antimicrobial peptides the number of positively charged residues is not a critical factor for activity and, following the model and SAR studies, the charge can be reduced from +7 in **R-22** to +5 in **P2-11**, or to +4 in **P3-10**. Furthermore, the character of the charged residue is also not critical, since mutations of Arg to Lys are very well tolerated.

The experiments performed with different concentrations of divalent cations, $Ca^{2+}$ and $Mg^{2+}$, revealed that these peptides are sensitive to the presence of these cations. This might indicate that prior to internalisation they interact with the negatively-charged cell membrane, and high concentrations of cations disturb this interaction. However, it is also known that high concentrations of these divalent cations in *in vitro* assays reduces the activity of RNase P.[199,200]

It is noteworthy that the most active peptidomimetic, **P2-05**, has an amphipathic character, having hydrophobic and hydrophilic faces (Figure 78), which makes it like many known antimicrobial peptides. The amphipathic character might mean that the RNase P peptidomimetics have a tendency to lyse red blood cells, however, this effect is fortunately not observed.

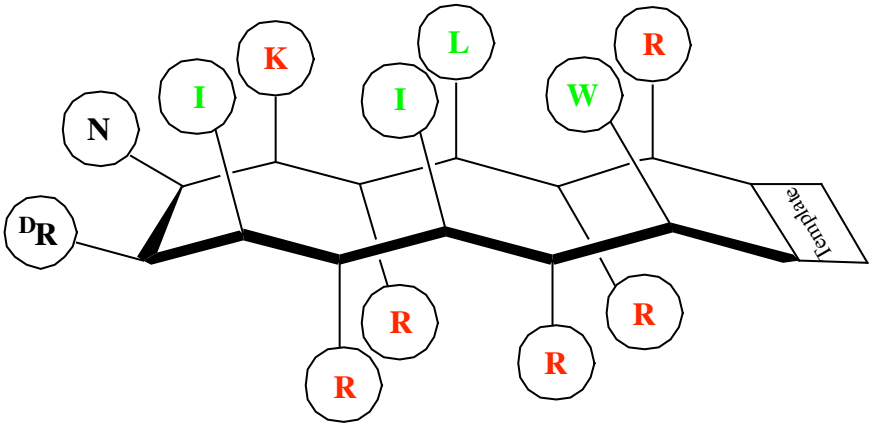



**Figure 78.** Topology of P2-05. Illustration of its amphipathic character, in red are polar amino acid and in green the aliphatic/aromatic amino acids.

The designed peptides are rich in arginine residues, which can make them susceptible to proteolytic degradation in biological settings. Selected peptides were tested in a plasma stability assay with human and mouse plasma, where they proved to be relatively stable under the experimental conditions. This stability might be attributable to the cyclic form of these peptides, which makes the peptide chain less accessible to the active sites of the proteolytic enzymes.

To conclude, it has proven possible to design potent and selective inhibitors of *E. coli* RNase P, which inhibit directly the processing of ptRNA, as detected by EMSA. Moreover, these molecules also inhibit the growth of several strains of *E. coli*, including clinically active pathogens. In addition, it was demonstrated that the peptides do not exhibit significant haemolytic activity and that they are stable in plasma, which makes them very attractive as potential future therapeutics. Further on-going studies should include optimisation of the lead sequences in order to increase activity and/or reduce the positive charge. Most important, further studies are required to prove that RNase P inhibition is really the mechanism of their antimicrobial activity.

# 5. EXPERIMENTAL PART

## 5.1. General notes

***Solvent and reagents.*** Reagents and solvents were purchased from *Aldrich*, *Acros*, *Merck* and *Fluka*. Ethyl acetate, diethyl ether, DCM, and cyclohexane were redistilled from $K_2CO_3$ or $CaCl_2$. Dry DCM and dry toluene were obtained by distillation from $CaH_2$. Other solvents or reagents were distilled, purified and/or dried using standard procedures.[201]

***Material and instrumentations.*** Flash chromatography was carried out on silica gel 60 (230-400mesh, 0.04-0.063mm) from *Merck* (Darmstadt, Germany) or *Chemie Uetikon AG* (Uetikon, Switzerland). Thin layer chromatography (TLC) was performed on silica gel 60 $F_{254}$ plates from *Merck*. Melting points (m.p.) were measured on a Kofler apparatus connected to a NIKON YS 100 microscope (Egg, Switzerland).

$^1H$ NMR and $^{13}C$ NMR we measured at 300 MHz and 75MHz, respectively, on a Bruker ARX-300 specrometer (Bruker, Rheinstetten, Germany). Chemical shifts are given in ppm relative to the internal tetramethylsilane (TMS) in organic solvents or sodium 3-trimethylsilyl(2,2,3,3-$D_4$)propionate (TSP) in $H_2O/D_2O$.

MS spectra were measured by the MS service of the University of Zurich. ESI mass spectra were recorded either on a Waters Q-tof Ultima API Instrument (Milford, USA), on a Bruker ESQUIRE-LC quadrupole ion trap instrument or on a Finnigam TSQ-7- instrument. MALDI MS spectra were recorded on Bruker Autoflex I MALDI-TOF. LC-MS analysis of the peptides was performed on a reverse-phase high performance liquid chromatography (RP-HPLC) system connected to a ThermoFinnegan Quadrupole-ESI Mass Spectrometer (Thermo Fisher Scientific, Waltham, USA).

Purification and analysis of peptides were carried out by RP-HPLC using a dual pump Äkta purifier (Amersham Pharmacia Biotech) system. The columns used were: Vydac$^{TM}$ GraceVydac 218TP54 (C18, 5μm, 4.6x250mm), Agilent Zorbax Eclipse XDB-C18 (5μm, 4.6x250mm) for analytical separations, Agilent Zorbax 300SB-C18 (7μm, 21.2x250mm) and Waters Xbridge$^{TM}$ Prep-C18 (5μm, 19x50mm) for

preparative separations. UV detection was performed at λ=226 nm and 278 nm. Solvents used were acetonitrile (*Acros Organics*) and distilled water, each containing 0.1% trifluoroacetic acid (TFA).

## 5.2. Peptide synthesis

***Solvents and amino acids.*** NMP and piperidine were purchased from *Acros Organics* (Geel, Belgium) and used without further purification. DMF was purchased from *Acros Organics* and redistilled under vacuum from ninhydrin. DIEA was purchased from *Acros Organics* and redistilled under vacuum first from ninhydrin and subsequently from KOH. HBTU was purchased from *Novabiochem* and HOBt, HATU and HOAt were purchased from *Fluka* and used without further purification. Amino acids were purchased from *Novabiochem*, *Bachem*, *NeoMPS*, and *Advanced ChemTech*. Resins: 2-chlorotritylchloride resin (100-200mesh) with a loading of 1.2mmol/g and Rink Amide MBHA resin with a loading of 0.69mmol/g were purchased from *Novabiochem*. The following N-α-Fmoc protected L/D-amino acids were used for routine Fmoc-SPPS:

Fmoc-Ala-OH, Fmoc-Asp(tBu)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Cys(Acm)-OH, Fmoc-Gly-OH, Fmoc-Glu(tBu)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Ile-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Lys(Alloc)-OH, Fmoc-Phe-OH, Fmoc-Pro-OH, Fmoc-Ser(tBu)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Trp(Trt)-OH, Fmoc-Val-OH.

Additionally some unnatural amino acids were used: cyclohexylalanine, Fmoc-Cha-OH, (*NeoMPS*); diaminobutyric acid, Fmoc-Dab(Boc)-OH, (*NeoMPS*); 2-L-aminobutyric acid, Fmoc-Abu-OH, (*NeoMPS*); β-phenyl-phenylalanine, Fmoc-Bpa-OH (*Sigma-Aldrich*); pipecolic acid, Fmoc-Pip-OH, (*NeoMPS*); biphenylalanine, Fmoc-Bip-OH, (*Advanced ChemTech*); homo-L-phenylalanine, Fmoc-Hfe-OH, (*Advanced ChemTech*); 3-chloro-L-phenylalanine, Fmoc-Phe(3-Cl)-OH, (*Advanced ChemTech);* 1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid, Fmoc-Tic-OH, (*NeoMPS*); 4-chloro-L-phenylalanine, Fmoc-Phe(4-Cl)-OH, (*NeoMPS*); ornithine, Fmoc-Orn(Boc)-OH (*NeoMPS*).

***Loading of amino acid on the resin.*** For the coupling of the first amino acid to 2-chlorotrityl chloride resin, a loading of 0.6 mmol of amino acid per gram of resin was typically used. The calculated amount of N-α-Fmoc protected amino acid was dissolved in dry DCM (10 ml per g of resin). To the pre-swollen resin (with dry DCM) the amino acid solution with 4 eq of DIEA was added. The reaction mixture was agitated for 120 min. The excess of unreacted resin was blocked by three consecutive washings with DCM/MeOH/DIEA (17:2:1) and then the resin was washed with dry DCM, DMF and DCM. The resin was dried overnight under vacuum over KOH. Estimation of the level of loading was performed by measuring the UV absorption at 290 nm after Fmoc-deprotection of ~1-2 mg dry resin in 25% piperidine/DMF for ~15 min. The blank sample contained only the 25% piperidine/DMF. The loading was calculated according to the equation:

$$\text{Loading (mmol/g)} = (\text{Abs}_{\text{sample}})/(\text{mg of resin} \times 1.75)$$ [202]

**Rink amide MBHA resin** was used directly in the SPPS procedure, with Fmoc deprotection as the first step of the synthesis.

***Peptide synthesis.***

Automated single peptide synthesis. Peptides were synthesised on an Applied Biosystems ABI433A automated synthesiser coupled to a Perkin Elmer UV/VIS detector using the FastMoc® chemistry protocol on a 0.25 mmol scale. In 0.25 mmol FastMoc chemistry, 4 eq of Fmoc-protected amino acid and 4 eq of coupling reagents: HBTU/HOBt (0.45 M in DMF), DIEA (2 M in NMP) are used for each coupling reaction. Deprotection is performed in 20% piperidine in NMP, monitored by UV and repeated if needed.

Automated parallel peptide synthesis. Synthesis of peptide libraries was carried out on a MultiSynTech Syro II peptide synthesiser on ~0.07 mmol scale using 4 eq of Fmoc-protected amino acid (0.3 M in DMF) and of coupling reagents HBTU/HOBt (0.45 M in DMF), DIEA (2 M in NMP). The coupling reactions were performed twice for 45 min. The deprotection was carried out twice in 40% piperidine/DMF for 5 min.

Manual coupling of amino acids. Coupling of non-standard amino acids was carried out manually with 2 eq of the Fmoc-protected amino acid activated with 1.9 eq of HATU/HOAt and 4 eq of DIEA in 10 ml dry DMF/g resin. The coupling

reactions were carried out for 4 h under vigorous agitation, and completion of the coupling reaction was monitored by the Kaiser test for primary amines[203] or chloranil test for secondary amines.[204]

***Cleavage from the resin.*** The linear, side chain-protected peptides assembled on the **2-chlorotrityl resin** were cleaved from the resin by treatment up to 10 times with 0.8% TFA/DCM for 2 min. The TFA filtrates were collected into separate flasks containing a few drops of DIEA to neutralise the acid and the solutions were evaporated under reduced pressure to yield crude linear fully protected peptides. The cleavage from the **rink amide MBHA resin** was carried out together with the final deprotection step using a cocktail of TFA:$H_2O$:TIPS [95:2.5:2.5 (v/v)] for 4 h. The filtrate was concentrated on a rotary evaporator and then the precipitation of crude fully-deprotected peptide was performed with ice-cold ether.

***Acetylation of N-terminal amine group.*** The acetylation of N-terminal amine groups was performed with peptides assembled on the rink amide MBHA resin. After the last Fmoc deprotection, the resin with fully-protected peptide was incubated twice with a solution of acetic anhydride (19 ml), DIEA (9 ml), and HOBt (0.8 g) in NMP (372 ml) (capping solution). The acetylation was monitored by a standard Kaiser test.

***Cyclisation and deprotection.*** Linear precursors (after cleavage from the chlorotrityl resin) were cyclised in DMF with 1.5 eq of HATU/HOAt and 6 eq of DIEA over 4 h. Concentrations of ~5 mg of crude precursor per ml DMF were used. After the reaction the mixture was concentrated under vacuum, yielding a yellow/brown oil containing a crude cyclic fully protected peptide. The final deprotection was performed in TFA/$H_2O$/TIPS (95:2.5:2.5) cocktail (10 ml) for 4 h. The solution was concentrated under vacuum and precipitation with ice-cold ether yielded a crude fully deprotected cyclic peptide.

***Disulfide bridge formation.*** The cysteines to be oxidised were protected with Acm protecting groups, which are stable to the standard final deprotection procedure. Pre-purified, fully deprotected (except for the Acm groups on Cys residues) cyclic peptides were dissolved in 4:1 AcOH:$H_2O$ mixture and treated with 10 eq of $I_2$ in methanol. The reaction was stirred for 30 min, then the excess of iodine was quenched

with a solution of ascorbic acid (50 mM) in citrate buffer (10 mM, pH=5.0). The reaction mixture was lyophilised. A dry crude product was redissolved in water and purified by RP-HPLC.

## 5.3. EMSA

***Run-Off Transcription of BIV TAR RNA.*** The synthetic oligodeoxynucleotide ("bottom strand") with a sequence complementary to that of BIV TAR, followed by a sequence complementary to the DNA-top-strand, 5'-GGC TCG GAG CTA ATG AGC TAC ACG CCT ATA GTG AGT CGT ATA A-3', was bought from *IDT-DNA technologies* (Iowa, USA) on a 1 μM scale. The DNA template provided was $OD_{260}$=11.1 (24.29 nmoles). The template was dissolved in sterile water (200 μl) and NaOAc (3 M, 35 μl) and absolute EtOH (600 μl) were added. The solution was kept at –20°C overnight and then centrifuged at 12'000 rpm, 4°C for 30 min. The DNA pellet was washed with 70% EtOH, centrifuged and dried. Then, the pallet was dissolved in sterile water, a $UV_{260}$ absorbance of a 100-fold diluted sample was measured, which gave $OD_{260}$=7.04 (15 mmol). The "bottom strand" oligonucleotide was dissolved in sterile water (1.875 ml) to give an 8 μM solution. The "top strand" (5'-TAA TAC GAC TCA CTA TAG-3', *IDT-DNA technologies*) was prepared in the same way to give an 8 μM stock solution. The BIV TAR RNA was prepared by run-off transcription. The run-off transcription was performed on a 100 μl scale in an Eppendorf tube. For this, sterile water (20 μl), "bottom strand" BIV TAR RNA (8 μM, 10 μl), "top strand" (8 μM, 10 μl), and $Mg_2Cl$ (1 M, 5.6 μl) were mixed in an Eppendorf tube, vortexed shortly and heated to 80°C for 2-3 min. The reaction mixture was left to cool down to rt for ~10 min, then NTPs (20 mM of each; 20 μl of a mixture of ATP, GTP, CTP, UTP (40 μl, 100 mM each, *Pharmacia*)), water (40 μl), PEG (20 μl, 8 K, 40%, *Axygen Biosciences*), [$\alpha$-$^{32}$P]-CTP (40 μl, 3000 Ci/mmol, *Amersham*), a transcription buffer (5 μl; 20x Tris-HCl (800 mM, pH=8.1), spermidine (20 mM, *Sigma-Aldrich*), Triton X-100 (0.2%, *Sigma*), and DTT (150 μl, 100 mM), were added and heated for 2-3 min at 37°C. T7 RNA polymerase (20 U/ml, 10 μl, *Ambion*) was added, the mixture was vortexed shortly and incubated at 37°C for 4h.

Afterwards, the transcription reaction was passed through a *Sephadex G-25* column (*Sigma-Aldrich*) to remove most of the unincorporated [$\alpha$-$^{32}$P]-CTP. To the eluted RNA, denaturing loading dye (TBE 2x, in 7 M urea, ~80 µl) was added and the mixture was loaded on a 20% denaturing polyacrylamide gel (20 cm x 52 cm) [75 ml of 40% polyacrylamide:bisacrylamide 19:1 (*Acros*), 63 g urea, 15 ml TBE 10x (890 mM Tris-base, 890 mM boric acid, 890 mM $Na_4EDTA$), 1500 µl APS (0.1 g/ml, *Acros*), 150 µl TEMED (*Acros*)]. The gel was run for ~24 h with 1x TBE (until the xylene cyanol dye in the loading dye had reached almost the bottom of the gel), at 900 V, 500 mA, 30 Watts (constant power). The gel was then removed (marked for reference with radioactive CTP and ink) and exposed on a phosphor image screen (*Molecular Dynamics*) for 30 min. The phosphor image screen was then scanned on a phosphor image scanner (*Molecular Dynamics*), and the autoradiogram was analysed with *Image Quant 5.0*. The wet gel was overlapped with the scanned autoradiogram (with the aid of the marker reference radioactive ink), and the RNA band was excised accurately. The main product band (N=28 mer) appeared at the expected height relative to the xylene cyanol dye (above the second, slower band of dye) indicating that the expected product (N):

5'-GGC UCG UGU AGC UCA UUA GCU CCG AGC C-3'

had been transcribed and was well separated from N+1 side product (above the intensive product band). The gel band was cut into small pieces and crashed gently into small pieces in a Falcon sterile tube and left mixing overnight with eluting buffer (1 ml, $H_2O$:NaOAc(3 M):$Na_2PO_4$ (pH=7.5, 0.5 M) (v/v 3:1:1)). The procedure was repeated 3 times, every time the eluting buffer was removed and a fresh portion of it was added to the pieces of the gel. All 3 portions of the elutant buffer were combined and passed through a 0.2 µl Millipore™ filter to remove any remaining gel pieces. The eluate was divided into Eppendorfs (400 µl in each), NaOAc (40 µl, 3 M) and absolute EtOH (900 µl) were added to each one and the solutions were kept at –20°C overnight for the RNA to precipitate, and then centrifuged at 12'000 rpm, at 4°C for 45 min. The solvent was decanted from above the pellet of the RNA, the pellet was washed with 70% EtOH, centrifuged again, solutions decanted again, dried on a high speed rotary vacuum and redissolved in sterile $H_2O$ (1 ml). The solution was dialysed against phosphate buffer (1 l, 10 mM, pH=6.6) overnight at 6°C. The dialyzed RNA was speed-vacced to dryness and dissolved again in sterile $H_2O$ (1 ml). A $UV_{260}$

absorbance was measured to calculate the final concentration of the BIV TAR RNA. The BIV TAR solution of 0.9 μM was prepared.

***Gel retardation assay.*** For each RNA-peptide binding reaction, a peptide at different dilutions was mixed in a sterile Eppendorf tube with 5 μl of assay buffer (100 μl contains: Tris (0.5 M, pH=8.0, 10 μl), KCl (0.5 M, 10 μl), DTT (1 M, 20 μl), tRNA (*E. coli*, 10 mg/ml, 0.28 μl), Triton X-100 (10%, 2 μl), sterile $H_2O$ (48 μl) and RNA stock solution (10 μl)). The mixture was incubated in ice for 30 min. Then, a native loading dye (5 μl, 10% glycerol, 0.1% bromophenol blue, 0.1% xylene cyanol) was added to each binding reaction, vortexed, centrifuged and loaded on a 0.4 mm, 12% native gel (prepared from sterile $H_2O$ (30 ml), acrylamide:bisacrylamide (29:1, 40%, 16.25 ml), TB [10x (890 mM Tris-base, 890 mM boric acid), 25 ml], Triton X-100 (10%, 0.5 ml), APS (0.1 g/ml, 0.5 ml) and TEMED (500 μl)). The gel should be left for at least 6 h to polymerise before electrophoresis. The native gels were run at 500 V, 150 mA, 15 Watts, at 4°C for 3.5 h. Then, they were dried at 70°C for 30 min, exposed on a phosphor image screen overnight and scanned on a phosphor scanner. The autoradiograms were analysed with *Image Quant 5.0*.

## 5.4. NMR analysis of peptides

***Data collection and analysis.*** 1D and 2D spectra were recorded at 600 MHz on a Brucker DRX-600 or Bruker AV-600 spectrometer, or at 700 MHz on a Bruker AV-700 spectrometer. Typically the spectra were recorded at a peptide concentration of 8 mg/ml or (if less) with the maximum material available. The solvent systems used were $H_2O/D_2O$ 9:1 pH=2.3 or $D_2O$ pH=2.3. The pH (uncorrected for isotope effect) for aqueous solutions was adjusted with 0.01 M solution of HCl or NaOH in $H_2O/D_2O$ 9:1. Water suppression was performed by presaturation using a power level of 5dB for the suppression. Spectral assignments were based on DQF-COSY, TOCSY and NOESY or ROESY spectra. Distance restraints were obtained from NOESY spectra with a mixing time of 250 ms. Spectra were typically collected with 1024x256 complex data points zero-filled prior to Fourier transformation to 2048x1024, and transformed with a cosine-bell weighting function. Data processing was carried out with TOPSPIN-NMR (Bruker) and XEASY.[205] The chemical shift values are given in

ppm relative to internal sodium 3-trimethylsilyl(2,2,3,3-$D_4$)propionate (TSP). The coupling constants *J* are given in Hz.

***Structure calculations.*** To derive NOE distance restraints, it was assumed that the initial rate approximation is valid and that each peptide rotates as a single isotropic rotor. The NOEs were determined from NOESY spectra measured with mixing times of 250 ms for all the peptides in $H_2O/D_2O$ 9:1 and $D_2O$, at temperatures of 298 and 278 K (to resolve CαH-XαH NOE cross-peaks close to the water signal). Cross-peak volumes were determined by integration using XEASY software. The relative cross-peak volumes were assumed to be proportional to $r^{-6}$ and were used to derive distance restraints. The structure calculations were performed using the program DYANA, which is based on restrained molecular dynamics in torsion angle space.[100] Starting from 100 randomised conformations a bundle of 20 conformations are selected, which have the lowest DYANA target energy function. The program MOLMOL was used for structural analysis and visualisation of the molecular models.[101]

***Hα chemical shifts relative to random coil chemical shifts.*** The differences between the CαH chemical shift values determined experimentally on the bases of $^1$H NMR recorded at 298K in $H_2O/D_2O$ 9:1 for each peptide and the random coil values were calculated following the protocol from Maynard and Griffiths-Jones.[102,103]

***H-D exchange experiments.*** The dry sample of the peptide (8-10 mg/ml, ~ 4.5 mM) was dissolved in a solution of $D_2O$. The experiments were carried out at 298 K. The first $^1$H NMR spectrum was recorded immediately after dissolving, and then the spectra were recorded every 3 min for the first hour, then after every 5 min for the second hour, then after every 15 min for the next 2 h. Then up to 24 spectra were recorded every one hour. If necessary measurements were repeated once a day until all the signals disappeared. The intensity of each signal was normalized relative to the corresponding value for a non-exchangeable peak for each data set. First order rate constants, were calculated following the protocol of Seebach et al.[206] {gradient of the plot: $\ln[I(NH_{exchangeable})/I(NH_{non\text{-}exchangeable})$ vs time }

***The temperature dependence of amide proton chemical shifts.*** $^1$H 1D and TOCSY NMR spectra were recorded for peptides at a concentration of 8-10 mg/ml in $H_2O/D_2O$ 9:1 (pH=2.3) using at least 6 different temperatures in the range of 278-328K.

## 5.5. Cell up-take studies

### 5.5.1. Materials

HeLa cells were obtained from the American Type Culture Collection ATCC (Rockville, MD). RPMI 1640 Medium with GlutaMax, foetal calf serum (FCS), trypsin-EDTA (10x), penicillin-streptomycin, sodium pyruvate (100mM), HEPES buffer (1M, pH 7.4) and Dulbecco's phosphate-buffered saline (D-PBS, pH 7.4) without calcium and magnesium were purchased from *Gibco*® (Invitrogen). Hoechst 33342, DAPI were from *Molecular Probes* (Leiden, The Netherlands). Dako fluorescent mounting medium was purchased from *DAKO Corp*. (Carpinteria, CA). 5(6)-Carboxyfluorescein was from *Fluka* (Buchs, Switzerland), and 5(6)-Carboxyrhodamine was obtained as Rhodamine WT from *Abbey Color* (Philadelphia, PA), as a 20% aq. sodium salt solution. This solution was acidified and the acid form was purified on reverse-phase HPLC. Cell culturing flasks (75 $cm^2$) and 24-well plates were from *TPP* (Trasadingen, Switzerland). Round coverslips (no.1.5) were purchased from *Hecht* (Glaswarenfabrik Karl Hecht, Germany) and microscope slides and squared cover glasses were obtained from *Menzel* (Braunschweig, Germany).

### 5.5.2. Cell culture

HeLa cells were cultured as exponentially growing subconfluent monolayers at 37°C under 5% $CO_2$. The cell culture was in 75 $cm^2$ culture flasks in RPMI 1640 medium (high glucose) supplemented with 10% heat-inactivated FCS, 1% penicillin/streptomycin, 1% HEPES (1 M, pH=7.4), and 1% sodium pyruvate (100 mM). The medium was exchanged two times per week. Exponentially growing HeLa cells were seeded at a constant density of about 5-8x$10^4$ cells/$cm^2$ on glass cover glasses (12 mm) in 24-well plates. For experiments, cells were used 24 h post-seeding.

### 5.5.3. Confocal laser scanning microscopy (CLSM)

HeLa cells were seeded on coverslips in 24-well plates as described above. The culture medium was discarded, and the cells were washed three times with D-PBS (pH 7.4). The cell monolayer was incubated for 30 min at 37°C with the peptides dissolved in the medium at the appropriate concentration. Subsequently, cells were washed three times with PBS and incubated with Hoechst 33342 for nuclear staining. Cells were washed three times with PBS and inspected immediately in a medium solution without any fixation. For the fixed cells, the protocol was the same except for the incubation with Hoechst 33342, which was omitted. The cells were fixed using a solution of 2% formaldehyde/3% sucrose (v/v) in PBS for 5 min at rt. After fixation, cells were incubated for 5 min with DAPI for nuclear staining, washed three times with PBS and mounted with the fluorescent mounting medium. The distribution of fluorescence was analysed using a high-resolution TCS-SP2 laser scanning confocal microscope (*Leica Microsystems*, Mannheim, Germany) with a 63×, 1.4 NA plan apochromatic lens using HeNe 594 nm, HeNe 543 nm, Ar 488 nm, and Ar UV 405 nm lasers. To avoid cross talk, emission signals were collected independently. Image processing was performed using IMARIS software (Bitplane AG, Zurich, Switzerland).

### 5.5.4. Peptide synthesis

A selection of fluorescently labelled peptides was prepared for the up-take studies (Table 46).

| **Peptide** | **Sequence** | Origin of the peptide |
|---|---|---|
| Fl-BIV-Tat | Fl-SGPRPRGTRGKGRRIRRK-$NH_2$ | BIV Tat protein |
| Rh-BIV-Tat | H-SGPRPRGTRGKGRRIRRK[Rh]-$NH_2$ | BIV Tat protein |
| Fl-HIV-Tat | Fl-GRKKRRQRRRPPQ-$NH_2$ | HIV Tat protein |
| Rh-BIV2$^h$ | cyclo(RVRTRGKRRIRV$^D$P[Rh]P) | BIV library |
| Fl-BIV2$^h$ | cyclo(RVRTRGKRRIRV$^D$P[Fl]P) | BIV library |
| Rh-L1-07$^h$ | cyclo(RTRTRGKRRIRV$^D$P[Rh]P) | L1 library |
| Fl-L1-07$^h$ | cyclo(RTRTRGKRRIRV$^D$P[Fl]P) | L1 library |
| Fl-Ac-BIV2$^h$ | cyclo(RVRTRGKRRIRV$^D$P[Fl-Ac]P) | BIV library |
| Fl-Ac-L1-07$^h$ | cyclo(RTRTRGKRRIRV$^D$P[Fl-Ac]P) | L1 library |

**Table 46.** Fluorescently labelled peptides synthesized for this study. Fl=5(6)-carboxyfluorescein, Rh=5(6)-carboxyrhodamine; Fl-Ac= 5(6)-carboxyfluorescein diacetate, $^h$=indication that the peptide contains the 4-hydrazino-D-Pro in the template.

- **Synthesis of Rh-BIV-Tat**

H-Ser$^{65}$-Gly-Pro-Arg-Pro-Arg-Gly-Thr-Arg-Gly-Lys-Gly-Arg-Arg-Ile-Arg-Arg-Lys$^{82}$(Rh)-$NH_2$

Rink amide resin (*Novabiochem*) (0.6 mmol/g) was used for the synthesis of the linear peptide. The chain assembly was performed using an ABI433A peptide synthesizer on a 0.25 mmol scale using the standard SPPS protocol described in Paragraph 5.2. For Lys82 Fmoc-Lys(Alloc)-OH was used. After completion of the synthesis, without removing the Fmoc-group on Ser65, the resin was washed with DCM and DMF. Alloc removal from Lys82 was carried out using $Pd(PPh_3)$ (1 eq), $PhSiH_3$ (60 eq) in dry DCM (15 ml) under argon for 3 h. After completion of the reaction, the resin was washed with DCM and DMF several times. 5(6)-Carboxy-rhodamine (1.5 eq) was coupled overnight manually to the free side chain amine group of Lys82 using HATU/HOAt (1.5 eq) and DIEA (6 eq) in dry DMF (20 ml).

The Fmoc-group of the N-terminus of the peptide was removed under standard conditions: 20% piperidine in DMF twice for 20 min. Cleavage of the peptide from the resin and deprotection was performed using standard conditions (cleavage cocktail TFA:$H_20$:TIPS, 95:2.5:2.5) for 4 h. The reaction mixture was then evaporated almost to dryness and the fully deprotected peptide was precipitated by addition of ice-cold ether. The resulting crude material was purified by PR-HPLC (Vydac 218TP1022 $C_{18}$ column, gradient 10-60% MeCN in $H_2O$ + 0.1%TFA) to give the fully deprotected linear peptide.

**Rh-BIV-Tat**: 12mg (6% yield), ESI-MS m/z [M+6H$^+$] 427.4.

- **Synthesis of Fl-BIV-Tat and Fl-HIV-Tat**

**Fl-BIV-Tat:** Fl-Ser$^{65}$-Gly-Pro-Arg-Pro-Arg-Gly-Thr-Arg-Gly-Lys-Gly-Arg-Arg-Ile-Arg-Arg-Lys$^{82}$-$NH_2$

**Fl-HIV-Tat:** Fl-Gly$^{48}$-Arg-Lys-Lys-Arg-Arg-Gln-Arg-Arg-Arg-Pro-Pro-Gln$^{60}$-$NH_2$

Rink amide resin (*Novabiochem*) (0.6 mmol/g loading) was used for the synthesis of the linear peptides. The chain assembly was performed using an

ABI433A peptide synthesizer on a 0.25 mmol scale using the standard SPPS protocol described in Paragraph 5.2. 5(6)-Carboxyfluorescein-*N*-succinimidyl ester (0.11 mmol, 1.3 eq, *Fluka*), was coupled overnight to the resin in DMF with DIEA (10 eq). A final cleavage from the resin was performed using a solution containing 95% TFA, 2.5% $H_2O$, 2.5% TIPS for 4 h at rt. Evaporation and precipitation by ice-cold ether gave the crude linear fully deprotected peptide. The peptides were purified by PR-HPLC (C18 column, Vydac 218TP1022, gradient 10-60% MeCN in $H_2O$ + 0.1%TFA) and analysed by ESI-MS or MALDI-MS.

**Fl-BIV-Tat**: 13mg, 6% yield, ESI-MS m/z $[M+6H^+]$=409.1

**Fl-HIV-Tat**: 68mg, 38% yield, MALDI-MS m/z $[M+H^+]$=2189

- **Synthesis of BIV2$^h$ and L1-07$^h$ (peptides containing 4-hydrazino-D-Pro)**

**BIV2$^h$**: cyclo(Arg-Val-Arg-Thr-Arg-Gly-Lys-Arg-Arg-Ile-Arg-Val-(4-$NH_2NH$)$^D$Pro-Pro)

**L1-07$^h$**: cyclo(Arg-Thr-Arg-Thr-Arg-Gly-Lys-Arg-Arg-Ile-Arg-Val-(4-$NH_2NH$)$^D$Pro-Pro)

Chlorotrityl resin preloaded with L-Pro (0.52 mmol/g) was used for the synthesis of both peptides. The resin (0.25 mmol) was first swelled with DMF and then treated with 20% piperidine in DMF to remove the Fmoc protecting group. The first two couplings were performed manually (2 eq HATU/HOAt, 6 eq DIEA in DMF) and monitored by the chloranil test. The further chain elongation was performed sequentially on an ABI433A with an appropriately protected amino acid (4 eq, extended coupling time), HBTU/HOBt for activation, DIEA as a base and NMP as a solvent. A solution of 20% piperidine in DMF was used for Fmoc-deprotection. After completion of the synthesis, the resin was washed with DCM and DMF and treated with 1% TFA in DCM (15 ml, 5 times). The filtrate was neutralized with DIEA and evaporated to dryness to yield the linear protected peptides. The peptides were then cyclized overnight at rt in DMF solution (10 mg/ml) with HATU/HOAt (3 eq) and DIEA (10 eq). After evaporation of the DMF, the products were dissolved in DCM, washed with 10% MeCN in water and dried to give the crude cyclic protected peptides. The cyclic protected peptides were treated with a standard final cleavage cocktail at rt for 4 h. Evaporation and precipitation with ice-cold ether gave crude cyclic deprotected peptides. The peptides were purified by RP-HPLC (Zorbax Eclipse

XDB-C18 column, gradient 10-60% MeCN in $H_2O$+ 0.1% TFA) and analysed by LC-MS:

**BIV2$^h$**: 155mg (35% yield), ESI-MS m/z ($M+4H^+$) 440.7.

**L1-07$^h$**: 145mg (33% yield), ESI-MS m/z ($M+4H^+$) 441.5.

- **Chemoselective labelling of BIV2$^h$ and L1-07$^h$ with 5(6)-carboxyfluorescein, 5(6)-carboxyfluorescein-diacetate, and 5(6)-carboxyrhodamine**

Each peptide BIV2$^h$ and L1-07$^h$ (~ 6 mg, 3.4 μmol) was dissolved in citrate buffer (0.5 ml, 50 mM, pH=5.0) and 5(6)-carboxyfluorescein-diacetate-succinimidyl ester (1.0 eq), 5(6)carboxyfluorescein-succinimidyl ester (1.0 eq) [both from *Fluka*] or 5(6)carboxyrhodamine-succinimidyl ester (1.0 eq) [freshly prepared with DIC and N-hydroxysuccinimide] dissolved in DMF (0.1 ml) was added to the solution. The reaction was stirred for 5 d at rt and its progress was followed by LC-MS. After completion the reaction mixture was frozen and lyophilised. The crude residue was dissolved in water and purification was performed on the RP-HPLC (semi-preparative Zorbax Eclipse, $C_{18}$, gradient 10-90%MeCN in $H_2O$+0.1%TFA) to give the desired product.

**Fl-BIV2$^h$**: 3.7 mg, 52% yield, MALDI m/z ($M+H^+$) 2118.

**Fl-L1-07$^h$**: 4.2 mg, 59% yield, MALDI m/z ($M+H^+$) 2120.

**Fl-Ac-BIV2$^h$**: 4.1 mg, 54% yield, MALDI m/z ($M+H^+$) 2201.

**Fl-Ac-L1-07$^h$**: 3.9 mg, 52% yield, MALDI m/z ($M+H^+$) 2203.

**Rh-BIV2$^h$**: 2.3 mg, 23% yield, MALDI m/z ($M+H^+$) 2228.

**Rh-L1-07$^h$**: 2.9 mg, 38% yield, MALDI m/z ($M+H^+$) 2230.

## 5.6. Bacteriological experiments

***Materials.*** For the preparation of all solutions and media, water was purified using an Elgastat® UHP-UF water purification system. Unless noted otherwise, all experiments with bacterial cultures were carried out in Difco™ Mueller Hinton broth (Becton Dickinson, Sparks, MD). For solid media Difco™ Mueller Hinton Agar (containing 17.0 g/l of agarose) was used. All media were prepared

according to the manufacturer's instructions and sterilised by autoclaving for 20 min at 121°C. The microorganisms used were ***Escherichia coli*** (ATCC 25922, 2138E, 2139E, 2140E, 2143E, 2144E, 3459E), ***Pseudomonas aeruginosa*** (ATCC 27853, PAO1), ***Staphylococcus aureus*** (ATCC 29213, ATCC 25923) and ***Bacillus subtilis*** (Nr. 3256).

***Bacterial cultures.*** Bacterial cultures were incubated at 37°C in conical flasks and agitated under orbital shaking at 200 rpm. Overnight bacterial cultures were inoculated with single colonies from agar plates freshly incubated overnight and used in turn to inoculate sterile Muller-Hinton broth, typically at ratio of ~1:100 (v/v). To estimate growth, the cell density was estimated photometrically via the turbidity of the bacterial cell suspensions at 600 nm with a UV/Vis spectrometer. For short term storage, bacterial strains were stored on Mueller Hinton agar plates at 4°C and were used to inoculate fresh agar plates weekly. For long term storage, bacterial cultures were stored in 30% glycerol solution at -80 °C.

***MIC determination.*** The determination of MICs was carried out according to a standard protocol[207] with modifications of Steinberg et. al.[208] to account for unspecific adsorption of AMPs onto reaction vessel surfaces. Lyophilised AMPs were dissolved in 0.01% AcOH at a concentration of 1 mg/ml. For routine MIC determination, colony material from agar plates incubated freshly overnight was taken with an inoculation loop and transferred into aqueous NaCl (3 ml of 0.9%). The turbidity of the bacterial suspension was adjusted to McFarland standard 0.5, and 100 µl of this bacterial suspension was transferred into Mueller Hinton broth (10 ml) supplement with 0.02% BSA. In 96-well microtitre plates, two-fold serial dilutions of AMPs were prepared in Mueller Hinton broth. The broth was supplemented with 0.02% BSA and the final volume of each dilution was 50 µl. Finally, 50 µl of the bacterial suspension were used to inoculate the microtitre plates to give a final cell density of approximately $5x10^5$ CFU/ml in a final volume of 100 µl in each well. Alternatively, for MIC determinations corresponding to higher cell density of the inoculum, overnight cultures of bacteria were diluted ~1:100 into fresh Mueller Hinton broth and incubated for 105 min under standard conditions to yield approximately $2x10^8$ CFU/ml (for *P. aeruginosa* PAO1). These bacterial suspensions were then used to inoculate the two-fold series of dilutions of test compounds, which

were prepared in Mueller Hinton broth supplemented with 0.04% BSA. The microtitre plates were covered and incubated overnight at 37°C in an orbital shaker at 200 rpm. The MIC is determined by visual inspection as the minimum concentration at which no visible growth occurs after 18 to 20 h of incubation and it is expressed in μg/ml.

***Haemolytic activity.*** The haemolytic activities of AMPs and a control compound (protegrin-1, PG-1) against human red blood cells (hRBCs) were tested by a photometric assay. Fresh hRBCs were washed with PBS and centrifuged three times for 10 min at 2000 g. Test compounds at concentrations of 100 $\mu$g/ml were incubated with 20% (v/v) of hRBCs (final erythrocyte concentration of ~$0.9 \times 10^9$/ml) at 37°C for one hour. The samples were centrifuged at 2000 g, the supernatant was diluted twenty-fold in PBS, and the optical density was measured at 540 nm. The values corresponding to 0% and 100% lysis were determined by incubation of hRBCs with PBS or 0.1% Triton X-100 in water, respectively. The $OD_{540nm}$ corresponding to 100% lysis typically ranges from 1.6 to 2.0 and the haemolytic activity obtained for the test compounds is expressed as a percentage relative to the 100% value.

## 5.7. Proteolytic stability

### 5.7.1. Trypsin digestion assay

***Materials.*** A standard procedure was conducted in Tris buffer (50 mM, pH=8.0) containing 10 mM $CaCl_2$ (assay buffer). Trypsin (from Bovine Pancrease, 13500 BAEE units/mg solid) was purchased from *Sigma-Aldrich*. BAPNA (benzoyl-Arg-p-nitroanilide) was purchased from *Fluka*. DL-Tryptophan, an internal control for HPLC integration, was obtained from *Fluka*. An analysis of the assay mixture was performed by analytical RP-HPLC (Agilent Zorbax Eclipse XDB-C18, 5 μm, 4.6x250 mm, gradient 10-50%, 6CV, MeCN in $H_2O$+0.1%TFA).

***Trypsin activity.*** Trypsin activity was tested by the method published first by F. Erlanger et al.[124] and modified by C.Temporini et al.[125] and L. Ocampo and J.M. Ezquerra.[209] BAPNA (43.5 mg) was dissolved in DMSO (1 ml) and the solution was

diluted to 100 ml with assay buffer to give a 1 mM stock solution. To BAPNA (2.5 ml, 1 mM) was added water (0.5 ml) and the mixture was allowed to equilibrate in a thermostatically controlled bath at 25°C. At zero time, 50 µl of trypsin solution (0.1 mg/ml in 1 mM HCl) was added, and the reaction was allowed to run for 1800 sec (the UV absorption at 410 nm was collected every second). A suitable control without enzyme was also used. The reaction was terminated by addition of 30% acetic acid (0.5 ml). The controls showed no self-hydrolysis of substrate. The BAPNA-unit (U) of trypsin activity is defined as the amount of enzyme that hydrolyses 1 µmol of substrate per minute at pH=8.0. The BAPNA-units of assayed trypsin were calculated by means of a calibration curves in which $UV_{410nm}$ was plotted against concentration of released *p*-nitroaniline (*p*-NA). The $\Delta A_{410nm}$/min from the fitted curves was used in following equation:

$$\mathbf{U\ (\mu mol/min)= \frac{\Delta A410nm/min \times 1000 \times reaction\ volume\ (ml)}{8800 \times mg\ of\ protein}}$$

where 8800 is the extinction coefficient of *p*-NA at 410nm. The average trypsin activity was found to be 0.4023 U.

***Trypsin digestion assay.*** A standard procedure was conducted at 37°C in Tris buffer. A solution of trypsin (1 mg/ml) in 1 mM HCl was prepared, and then diluted 10x with the assay buffer. A typical experiment was performed by adding of trypsin solution (20 µl, 0.1 mg/ml in the assay buffer) to a solution of peptide (300 µl, 1mg/ml in $H_2O$), followed by addition of tryptophan (300 µl, 0.5 mM in the assay buffer) and the assay buffer (380 µl) to obtain a final reaction volume of 1 ml. At the time points (typically 0, 1, 2, 5, 10, 15, 30, 60, 120, 240 min) 100 µl of the reaction mixture was removed and the trypsin was inactivated by addition of 30% AcOH/$H_2O$ solution (50 µl). Samples were analysed by RP-HPLC with integration of the peptide and tryptophan peaks.

### 5.7.2. Plasma stability studies

***Materials.*** Human plasma was obtained from Basler Blutspendedienst, and mouse plasma from Harlan Sera Lab, UK. The 1 mM peptide solutions were prepared from lyophilised samples in water. Assays were performed at 37°C in PBS buffer, pH=7.4. To terminate the assay, the assay solution was transferred to Millipore Filtration

Plates (MSRPN0410) filled with acetonitrile (HPLC-grade, +2% Formic Acid) and then filtered. For this, the Millipore MultiScreen HTS Vacuum Manifold (MSVMHTS00), Deep Well Collar (MSVMHTS0D) and Millipore Receiver Plates PP (MDCPN2M50) were used.

***Plasma stability assay.*** The peptide solution (35 $\mu$l, 0.1 mM in PBS) was added to freshly thawed plasma (315 µl), and incubated at 37°C. At five time points (0, 15, 30, 60, 120, 240 min), 50 µl of the reaction mixture was transferred into filtration plate wells containing 150 µl of MeCN. After 2 min of shaking (700 rpm) on a microplate shaker, the plate was placed on a Millipore receiver plate and a vacuum of ~-50 cmHg was applied until the filtration was completed. 100 µl of the filtrate was transfered to the microtiter plate and was subjected to LC-MS analysis. The assays were completed in triplicate.

***Assay assessment.*** The peak area counts from LC-MS analysis are used for the calculations. The triplicate values were averaged, unless obvious deviations were observed (and disregarded from calculations). The stability is expressed as a percent remaining after the initial time point 0 min (100% x $T_x/T_0$) and plotted graphically over time. Using Excel, the data are used to determine $t_{1/2}$ and a T240 min parameter (the fitted amount of peptide remaining after 240 min).

## 5.8. Circular dichroism spectra measurements

CD spectra were recorded on a Jasco J-715 spectropolarimeter (Jasco, Tokyo, Japan) at 25°C. A cylindrical, thermostated, quartz cell with 0.1 cm path length was used for measurements. The peptides were dissolved in PBS (10 mM, pH=7.4) to obtain a peptide concentration of approximately 20-30 µM. For further investigations, accurate concentrations were determined by amino acid analysis. Far-UV CD experiments were carried out in the wavelength range of 185 to 250 nm. Five scans were averaged for each sample, and the blank was subtracted from all spectra. The sample solutions were stored at rt for 30 min before measurements. The results are expressed as the mean residue ellipticity ($\theta$) with units of degrees $cm^2$ $dmol^{-1}$. Other experimental settings were 5 nm/min scan speed, 2.0 nm band width, and 2 sec response time.

## 5.9. Synthesis of Fmoc-*N*,*N*'-di-Boc-*L*-homoarginine (Fmoc-hArg(Boc)$_2$-OH)

***N, N'-Bis(tert-butoxycarbonyl)-guanidine (18).*** 1,4-Dioxane (50 ml) was added to a solution of guanidine hydrochloride (2.39 g, 25 mmol) and sodium hydroxide (4.0 g, 0.1 mol) in water (25 ml) and the resulting mixture was cooled to 0°C. Di-*tert*-butyl pyrocarbonate (12.0 g, 55 mmol) was added in one portion and the reaction mixture was allowed to warm to rt within 2 h. After stirring for 14 h (overnight) the mixture was concentrated *in vacuo* to one third of its original volume. The resulting suspension was diluted with water (50 ml) and extracted three times with ethyl acetate. The combined extracts were washed with 10% citric acid, water and brine, and dried over magnesium sulfate. After filtering and removal of the solvent under reduced pressure the crude product was purified by flash chromatography on silica gel (eluent: first DCM 100%, then DCM:MeOH 97:3). N, N'-Bis(tert-butoxycarbonyl)-guanidine (4.2 g, 64%) was obtained as a colourless powder, m.p.=139-141°C (lit. m.p.144[97]); $R_f$=0.34 (DCM:MeOH 97:3); $^1$H NMR (300 MHz, DMSO-$d_6$) 10.5-10.0 (broad s, 1H), 8.5-8.3 (broad s, 2H), 1.39 (s, 18H); ESI-MS: m/z [M+Na$^+$] 282.1.

***N,N'-Di-Boc-N''-trifluoromethanesulfonylguanidine (19).*** A solution of *N*, *N'*-bis(*tert*-butoxycarbonyl)-guanidine, (**18**) (1.8 g, 6.95 mmol) and triethylamine (1.16 ml, 6.8 mmol) in anhydrous DCM (25 ml) was cooled to –78°C under an atmosphere of Ar. Triflic anhydride (1.12 ml, 6.8 mmol) was added dropwise at such a rate that the reaction temperature did not exceed –65°C. When the addition was complete, the mixture was allowed to warm to rt within 4 h. The solution was washed with 1M sodium bicarbonate and water, and dried with magnesium sulfate. After filtering and removal of the solvent under reduced pressure the crude product was purified by flash chromatography on silica gel (eluent: DCM 100%). *N*,*N'*-di-Boc-*N''*-Trifluoromethanesulfonylguanidine (1.74 g, 65%) was obtained as a white powder, m.p.=117-119°C (lit. m.p.=115[97]); $R_f$=0.71 (DCM 100%); $^1$H NMR (300 MHz, DMSO-$d_6$) 11.0 (broad s, 1H), 1.4 (s, 18H); $^{13}$C NMR (75 MHz, $CHCl_3$) 154.3, 152.7 (C=O urethane, Boc), 83.2 ($\underline{C}(CH_3)_3$), 27.4 ($C(\underline{C}H_3)_3$). ESI-MS: m/z [M+Na$^+$] 314.1.

***Fmoc-N,N'-di-Boc-homoarginine (20).*** Fmoc protected lysine (393 mg, 1.07 mmol) was suspended in anhydrous DCM (25 ml) under Ar. To the amino acid suspension triethylamine (178 μl, 1.28 mmol) was added followed by addition of *N,N'*-di-Boc-*N''*-trifluoromethanesulfonylguanidine (500 mg, 1.28 mmol). The reaction was stirred overnight at rt, then another portion of Fmoc-Lys-OH (100 mg) was added to ensure completeness of the reaction. After 12 h, the reaction mixture was washed with 1M citrate buffer (pH=5.0) and brine, and dried over magnesium sulfate. After filtering and removal of the solvent under reduced pressure the crude product was purified by flash chromatography on silica gel (eluenting first with AcOEt:Hexane 1:1 + 1% AcOH, then 99.5% AcOEt+0.5% AcOH). Fmoc-*N,N'*-di-Boc-homoarginine was obtained as a white powder, m.p.=79-83°C (decomposition) (lit. m.p.=155[210]); $R_f$=0.23 (99.5% AcOEt + 0.5% AcOH); $\alpha^{25}_D$= +2.0 (*c* 0.04, MeOH); $^1$H NMR (300 MHz, $CDCl_3$) 8.80 (broad *s*, 1H, NH), 7.85 (*d*, *J* 7.8 Hz, 2H, Fmoc), 7.69 (*d*, *J* 7.8 Hz, 2H, Fmoc), 7.38 (*t*, *J* 7.4 Hz, 2H, Fmoc), 7.30 (*t*, *J* 7.4 Hz, 2H, Fmoc), 5.60 (broad *d*, *J* 8.2 Hz, 1H, NH), 4.40 (*d*, *J* 7.9 Hz, 2H, -$OCH_2$-(Fmoc)), 4.35 (*m*, 1H, CHα), 4.25 (*t*, *J* 7.9 Hz, 1H, CH(Fmoc)), 3.30-3.45 (*m*, 2H, $CH_2$(ε)) 1.85-1.98 (*m*, 2H, $CH_2$(β)), 1.65 (*m*, 2H, $CH_2$(δ)), 1.49 (*s*, 3H, $CH_3$), 1.47 (*s*, 3H, $CH_3$), 1.45-1.55 (*m*, 2H, $CH_2$(γ));); $^{13}$C NMR (75MHz, $CHCl_3$) 176.2 (C=O acid), 163.2 (C=N), 157.2 155.3, 152.7 (C=O urethane, Boc, Fmoc), 144.6, 142.2, 128.6, 128.0, 126.0, 120.8 (aromat, Fmoc) 84.2, 80.7 ($\underline{C}(CH_3)_3$), 67.0 (-$OCH_2$-(Fmoc), 50.5 (CH(Fmoc), 40.5 ($C_\alpha$), 33.2 ($C_\varepsilon$), 31.2 ($C_\delta$), 29.8 ($C_\gamma$), 29.1-28.9 ($C(\underline{C}H_3)_3$), 22.5 ($C_\beta$). ESI-MS: m/z [$M+H^+$] 611.7.

## 5.10. Synthesis of (2*R*,4*S*)-4-[*N,N,N'*-tris-(*tert*-butyloxycarbonyl)]-hydrazine-1-[(9*H*-fluoren-9-yl)-metoxycarbonyl]-proline

***(2R,4R)-N-Benzyloxycarbonyl-4-hydroxyproline (21).*** To a cooled solution (ice/water bath) of (2*R*,4*R*)-hydroxyproline (10 g, 60.4 mmol) in 1M NaOH (70 ml) and THF (60 ml), a solution of benzyl chloroformate (1.25 eq, 11.4 ml, 75.5 mmol) in THF (60 ml) was added, followed by addition of 1M NaOH (70 ml). The reaction was allowed to reach rt and was stirred overnight. The reaction mixture was diluted with water and 1M NaOH (~100 ml) was added to pH=12. The aqueous layer was washed with diethyl ether (3 x 50 ml), acidified to pH~3 with concentrated HCl, and extracted

with ethyl acetate (3 x 75 ml). The combined organic layers were washed with water (40 ml), brine (40 ml) and dried over $Na_2SO_4$. The solvent was removed *in vacuo* to give a colourless oil which precipitated from ether/hexane to give a white powder (12.1 g, 76%). m.p.=107°C (lit. 110.5-111.5 °C[211]), $R_f$=0.1 (AcOEt/Hexane 1:1); $\alpha^{25}_D$=+42.1° (*c* 0.05, MeOH) (lit. $[\alpha]_D$=+26.3[211]); $^1$H NMR (300 MHz, $D_2O$) 7.3-7.45 (5H, aromat), 5.12 (*s*, 1H, -$OCH_2$-), 5.08 (*d*, *J* 6.2 Hz, 1H, -$OCH_2$-) 4.4-4.5 (*m*, 2H, CH(α), CH(γ)), 3.55-3.7 (*dd*, $J_1$ 5.0 Hz, $J_2$ 11.0 Hz, 1H, CH(δ)), 3.32-3.45 (*t*, *J* 12.0 Hz, 1H, CH(δ)), 2.35-2.55 (*m*, *J* 5.0 Hz, 1H, CH(β)), 2.05-2.15 (*d*, *J* 12.0 Hz, 1H, CH(β)); $^{13}$C-NMR (75 MHz, $D_2O$): 176.0 (C=O acid), 155.9 (C=O urethane), 135.9, 128.6, 128.2, 127.5 (aromat), 69.0 (C(γ)), 67.5 (benzyl), 57.9 (C(α)), 54.0 (C(δ)), 37.3 (C(β)). ESI-MS: m/z [M+$H^+$] 266.2.

***(2R,4R)-N-Benzyloxycarbonyl-4-hydroxyproline benzyl ester (22).*** To a solution of (2*R*,4*R*)-*N*-benzyloxycarbonyl-4-hydroxyproline (21) (12 g, 45.2 mmol) in DMF, anhydrous potassium carbonate (2 eq, 12.47 g, 135.6 mmol), sodium iodine (0.1 eq, 0.72 g, 4.6 mmol) and benzyl bromide (3 eq, 16.26 ml, 135.6 mmol) were added under nitrogen. The mixture was stirred overnight at rt. The reaction mixture was diluted with water (100 ml) and extracted with ethyl acetate (3x60 ml). The organic layers were combined and washed with water (50 ml), brine (50 ml) and dried over $Na_2SO_4$. The solvent was removed *in vacuo* and the crude product was purified by flash chromatography (30% AcOEt/Hexane) to give the desired product as a colourless oil (15.7 g, 98%), $R_f$=0.32 (AcOEt/Hexane 1:1); $\alpha^{25}_D$=+19.8° (*c* 0.01, MeOH) (lit. $[\alpha]_D$=+34.2 (0.5 MeOH)[212]); $^1$H NMR (300 MHz, $CHCl_3$) 7.25-7.45 (10H, aromat), 5.0-5.3 (*m*, 4H, 2x -$OCH_2$Ph), 4.4-4.52 (*dd*, $J_1$ 19.1 Hz, $J_2$ 10.9 Hz, 1H, CH(γ), 4.35 (broad *s*, 1H, (CH(α)), 3.72 (t, *J* 12.7, 1H, CH(δ)), 3.55-3.68 (*td*, $J_1$ 3.6 Hz, $J_2$ 12.7 Hz, 1H, CH(δ)), 2.30-2.55 (*m*, 1H, CH(β)), 2.05-2.12 (*dd*, $J_1$ 8.1 Hz, $J_2$ 12.7 Hz, 1H, CH(β)); $^{13}$C-NMR(75MHz, $CDCl_3$) 172.9 (C=O ester), 154.9 (C=O urethane), 128.0, 128.2, 128.3, 128.6, 128.8, 135.5, 135.8 (aromat), 69.3 ($C_\gamma$), 68.9, 67.0 (-O$\underline{C}H_2$Ph), 58.0 ($C_\alpha$), 54.8 ($C_\delta$), 38.3 (Cβ). ESI-MS: m/z [M+$Na^+$] 378.1.

***(2R,4R)-N-Benzyloxycarbonyl-4-O-(p-nitrophenylsulfonyl)-proline benzyl ester (23).*** To a cooled solution (ice/water bath) of (2*R*,4*R*)-*N*-Benzyloxycarbonyl-4-hydroxyproline benzyl ester (17.7 g, 50 mmol) in DCM (150ml), NosCl (1.3 eq, 14.3

g, 65 mmol) and $Et_3N$ (1.5 eq, 10.4 ml, 75mmol) were added under nitrogen. The reaction was allowed to reach rt and was stirred overnight. The orange solution was extracted with 1M HCl (3 x 50 ml) and with saturated $NaHCO_3$, (3 x 50 ml). The organic layers were combined and washed with water (50 ml), brine (50 ml) and dried over $Na_2SO_4$. The solvent was removed *in vacuo* and the crude product was purified by flash chromatography (AcOEt/Hexane 1:2) to give the desired product as a pale yellow oil (25.7 g, 96%). $R_f$=0.4 (AcOEt/Hexane 40:60); $\alpha^{25}_D$ =+26.3° (*c* 0.1, $CHCl_3$); $^1$H NMR (300 MHz, $CHCl_3$) 8.25-8.35 (*t*, *J* 9.0, 2H, Nos) 7.95-8.05 (*dd*, $J_1$ 3.6, $J_2$ 9.0, 2H, Nos) 7.2-7.4 (10H, aromat), 5.05-5.35 (*m*, 5H, 2x -$OCH_2$-, CH(γ)), 4.55-4.7 (*ddd*, $J_1$ 2.7 Hz, $J_2$ 8.2 Hz, $J_3$ 25.5 Hz, 1H, CH(α)), 3.75-3.85 (*m*, 2H, CH(δ)), 2.4-2.6 (*m*, 2H, CH(β)); $^{13}$C NMR (75MHz, $CHCl_3$) 171.5 (C=O ester) 154.6 (C=O urethane), 143.0, 135.9, 128.8, 128.5, 128.4, 127.9, 124.4 (aromat), 79.2 ($C_\gamma$) 67.5, 67.1 (-O<u>C</u>$H_2$Ph), 57.6 ($C_\alpha$), 52.1 ($C_\delta$), 37.1 ($C_\beta$). ESI-MS: m/z [M+$Na^+$] 563.2.

***(2R,4S)-N-Benzyloxycarbonyl-4-[N-(tert-butyloxycarbonyl)hydrazino]-proline benzyl ester (24).*** To a solution of (2*R*,4*R*)-*N*-benzyloxycarbonyl-4-*O*-(*p*-nitrophenylsulfonyl)-proline benzyl ester (25.4 g, 47 mmol) in dioxane (180 ml), *tert*-butylcarbazate (8 eq, 49.7 g, 376 mmol) was added. The reaction was refluxed for 40 h. The solution was diluted with water (300 ml) and extracted with DCM (5 x 80 ml). The organic layers were combined and washed with water (50 ml), brine (50 ml) and dried over $Na_2SO_4$. The solvent was removed *in vacuo* and the crude product was purified by flash chromatography (eluting first with AcOEt/Hexane 1:3, then AcOEt/Hexane 1:2) as a preliminary purification and then by Kugelrohr distillation, (80°C, 0.1 atm) to give the desired product as a pale yellow oil (25.7 g, 96%). $R_f$=0.53 (AcOEt/Hexane 3:4); $\alpha^{25}_D$ =+24.9° (*c* 0.1, MeOH); $^1$H NMR (300 MHz, $CHCl_3$) 7.1-7.35 (10H, aromatic), 6.1-6.3 (broad *d*, *J* 31.8 Hz, 1H, NH), 5.05-5.2 (*m*, 2H,-$OCH_2$-), 4.95-5.05 (*d*, 2H, -$OCH_2$-), 4.35-4.55 (*m*, *J* 10.9 Hz, 1H, CH(γ)), 3.7-3.85 (*m*, 2H, CH(α), NH), 3.4-3.6 (*m*, 2H, CH(δ)), 2.1-2.2 (*m*, 1H, CH(β)), 1.9-2.0 (*m*, 1H, C(β)), 1.4-1.5 (3x *s*, 9H, 3x $CH_3$, Boc); $^{13}$C NMR (75 MHz, $CHCl_3$) 172.4 (C=O ester), 156.8 (C=O urethane, Boc), 154.9, (C=O urethane, Bn), 136.5, 135.5, 128.4, 128.3, 128.1, 128.0, 127.8 (aromat), 80.9 (<u>C</u>$(CH_3)_3$), 67.0, 66.8 (-O<u>C</u>$H_2$Ph), 58.2 ($C_\alpha$), 57.9 ($C_\gamma$), 50.9 ($C_\delta$), 34.9 ($C_\beta$), 28.2 (C(<u>C</u>$H_3)_3$). ESI-MS: m/z [M+$Na^+$] 492.2.

***(2R,4S)-N-Benzyloxycarbonyl-4-[N,N,N'-tris-(tert-butyloxycarbonyl)-hydrazino]-proline benzyl ester (25).*** To a solution of (2*R*,4*S*)-*N*-benzyloxycarbonyl-4-[*N*-(*tert*-butyloxycarbonyl)-hydrazino]-proline benzyl ester (7 g, 15 mmol) in DCM (100ml), di-*tert*-butyl dicarbonate (5 eq, 16.4 g, 74.5 mmol), DMAP (0.5 eq, 0.91 g, 74.5 mmol) and $Et_3N$ (5 eq, 4.7 ml, 74.5 mmol) were added. The reaction was stirred at rt for 4 h and its progress was followed by TLC. The solvent from the reaction mixture was removed *in vacuo* and the crude product was purified by flash chromatography (AcOEt/Hexane 1:3) to give the desired product as a yellowish oil (6.49g, 62%). $R_f$=0.34 (AcOEt/Hexane 1:3); $\alpha^{25}_D$=+32.0° (*c* 0.05, MeOH); $^1$H NMR (300 MHz, $CHCl_3$) 7.2-7.5 (10H, aromatic), 5.0-5.3 (*m*, 2H, 2x -$OCH_2$-), 4.65-4.8 (*m*, *J* 7.2 Hz, 1H, CH(α)), 4.4-4.65 (*ddd*, $J_1$ 3.6, Hz, $J_2$ 9.1, Hz, $J_3$ 19.0 Hz, 1H, CH(γ)), 3.8-3.9 (*m*, 1H, CH(δ)), 3.45-4.65 (*m*, 1H, CH(δ)), 2.52-2.35 (*m*, 1H, CH(β)), 2.08-2.15 (*m*, 1H, C(β)) 1.3-1.4 (27H, 9x $CH_3$, Boc); $^{13}$C NMR (75 MHz, $CHCl_3$) 173.5 (C=O, ester), 157.0 (C=O urethane, Boc), 153.0 (C=O urethane, Bn), 152.0 (C=O urethane Boc), 136.5, 135.5, 128.5-127.7 (aromat), 83.6, 81.7 ($\underline{C}(CH_3)_3$), 67.0, 66.9 (-O$\underline{C}H_2$Ph), 57.6 ($C_\alpha$), 56.3 ($C_\gamma$), 47.5 ($C_\delta$), 33.2 ($C_\beta$), 28.0-27.8 (C($\underline{C}H_3)_3$). ESI-MS: m/z [M+$Na^+$] 692.3.

(2*R*,4*R*)-*N*-Benzyloxycarbonyl-4-[*N*,*N'*-bis-(*tert*-butyloxycarbonyl)-hydrazino-proline was isolated as a minor product (2.6g, 31%).

***(2R,4S)-4-[N,N,N'-tris-(tert-butyloxycarbonyl)hydrazino]-proline (26).*** To a solution of (2*R*,4*S*)-*N*-benzyloxycarbonyl-4-[*N*,*N*,*N'*-tris-(*tert*-butyloxycarbonyl)-hydrazino]-proline benzyl ester (19.7 g, 29 mmol) in MeOH (300 ml), a spatula of 10% Pd/C (~200 mg) was added. The mixture was stirred for 30 h under hydrogen. The catalyst was removed by filtration through celite and the solvent was removed *in vacuo*. The crude product was precipitated from AcOEt to give a white powder (9.7 g, 75%). m.p.=204°C; $R_f$=0.185 (DCM:MeOH 9:1); $\alpha^{25}_D$ =+9.2° (*c* 0.09, MeOH); $^1$H NMR (300 MHz, $CHCl_3$) 4.45-4.55 (*t*, *J* 7.3 Hz, 1H, CH(γ)), 4.32-4.4 (*t*, *J* 7.7 Hz, 1H, CH(α)), 3.8-3.9 (*dd*, $J_1$ 7.4 Hz, $J_2$ 11.8 Hz, 1H, CH(δ)), 3.45-3.55 (*dd*, $J_1$ 7.4 Hz, $J_2$ 12.0, 1H, CH(δ)) 2.4-2.5 (*m*, 2H, CH(β)) 1.35-1.55 (27H, 9x $CH_3$); $^{13}$C NMR (75 MHz, $CHCl_3$) 174.8 (C=O acid), 155.3, 155.2, 155.0 (C=O urethane, Boc), 84.2, 82.6 ($\underline{C}(CH_3)_3$), 59.6 ($C_\alpha$), 57.3 ($C_\gamma$), 47.1 ($C_\delta$), 32.8 ($C_\beta$), 28.0-27.8 (C($\underline{C}H_3)_3$). ESI-MS: m/z [M+$H^+$] 446.2.

***(2R,4S)-4-[N,N,N'-tris-(tert-butyloxycarbonyl)-hydrazine]-1-[(9H-fluoren-9-yl)-metoxycarbonyl]-proline (27).*** To a solution of (2*R*,4*S*)-4-[*N*,*N*,*N'*-tris-(*tert*-butyloxycarbonyl-)hydrazino]-proline, (9 g, 20 mmol) in DCM (150 ml), Fmoc-*O*-succinimide (1.3 eq, 8.8 g, 26 mmol) and DIEA (2.2 eq, 7.2 ml, 44 mmol) were added. The reaction mixture was stirred overnight at rt. The solvent was removed *in vacuo* and the crude product was purified by flash chromatography (AcOEt/Hexane/AcOH 2:1/1%) to give a white powder (12.8 g, 72%). m.p.=89°C; $R_f$=0.46 (AcOEt/Hexane/AcOH 2:1/1%); $\alpha^{25}{}_D$=+23.3° (*c* 0.07, MeOH); $^1$H NMR (300 MHz, MeOH) 7.70-7.78 (*t*, *J* 7.3 Hz, 2H, aromatic), 7.54-7.62 (*t*, *J* 7.3 Hz, 2H, aromatic), 7.29-7.37 (*t*, *J* 7.5 Hz, 2H, aromatic), 7.22-7.28 (*t*, *J* 7.5 Hz, 2H, aromatic), 4.65-4.75 (*m*, 1H, CH(γ)), 4.1-4.5 (*m*, 4H, CH(α),-$OCH_2$-(Fmoc), CH(Fmoc)), 3.55-3.85 (*m*, 2H, CH(δ)), 2.4-2.6 (*m*, 1H, CH(β)), 2.15-2.35 (*m*, 1H, CH(β)) 1.4-1.6 (27H, 9x $CH_3$); $^{13}$C NMR (75 MHz, $CHCl_3$) 174.3 (C=O acid), 155.3 (C=O urethane, Fmoc), 153.1, 150.7 (C=O urethane, Boc), 143.6, 141.1, 127.6, 127.0, 125.0, 119.9 (aromat, Fmoc) 84.0, 82.0 ($\underline{C}(CH_3)_3$), 68.2 (-$OCH_2$-(Fmoc), 57.8 ($C_\alpha$), 56.4 ($C_\gamma$), 48.2 ($C_\delta$), 47.0 (CH(Fmoc)), 31.2 ($C_\beta$), 28.0-27.8 ($C(\underline{C}H_3)_3$). ESI-MS: m/z [M+$Na^+$] 690.3.

# 6. APPENDICES

## Appendix 1: Abbreviations

**AIDS** ……..…...acquired immunodeficiency syndrome
**AMP** ……...…...…….……..antimicrobial peptide
**AMPM** ……...……...…….antimicrobial peptidomimetics
**APS** .....……ammonium persulfate
**ARM** ……..…arginine rich motif
**BIV** ……bovine immunodeficiency virus
**BSA** .....bovine serum albumin
**Boc** ……*tert*-butyloxycarbonyl
**CD** ……..circular dichroism
**CDK** ……cyclin dependent kinase
**CFU** ……...…….colony forming units
**CTD** ……..carboxy terminal domain
**DCM** ……dichloromethane
**DIEA** ……...…….diisopropylethylamine
**DMF** ……..……N,N-dimethylformamide
**DMSO** ……..……dimethylsulfoxide
**DNA** ……deoxyribonucleic acid
**DQF-COSY** ……double quantum filtered correlated spectroscopy
**dsRNA** ……double stranded RNA-binding domain
**Fl** ……...……fluorescein
**EMSA** ……electrophoresis mobility shift assay
**eq** ……equivalent
**ESI** ……..……electrospray ionisation
**Fmoc** ……..……N-9H-fluoren-2-ylmethoxycarbonyl
**Fmoc-OSu** ……*N*-(9H-fluoren-2-ylmethoxycarbonyl)succinimide
**HATU** …….*O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*',*N*'-tetramethyluronium hexafluorophosphate
**HBTU** ……*O*-(benzotriazol-1-yl)-*N*,*N*,*N*',*N*'-tetramethyluronium hexafluorophosphate
**HEPES** ……4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid
**HIV** ……human immunodeficiency virus
**hnRNPs** ……heterogeneous ribonucleoprotein particles
**hnRNA** ……heterogeneous nuclear RNA
**HOAt** ……1-hydroxy-7-azabenzotriazole

**HOBt** ……1-hydroxybenzotriazole
**HPLC** ……high performance liquid chromatography
**hRBC** ……human red blood corpuscles
**$K_d$** ……dissociation constant
**KH** ……K-homology domain
**LC-MS** ……liquid chromatography-mass spectroscopy
**LTR** ……long terminal repeat
**MALDI** ……matrix-assisted laser desorption/ionisation
**MIC** ……minimal inhibitory concentration
**MH broth/agar** ……Mueller-Hinton broth/agar
**m.p.** ……melting point
**mRNA** ……messenger RNA
**MS** ……mass spectroscopy
**NHS** ……*N*-hydroxysuccinimide
**NMP** ……*N*-methylpyrrolidinone
**NMR** ……nuclear magnetic resonance
**NES** ……nuclear export signal
**NOESY** ……nuclear Overhauser enhancement spectroscopy
**NTPs** ……nucleotide triphosphate
**Pbf** ……pentamethyldihydrobenzofuran-5-fulfonyl
**PDB** ……protein data base
**PEG** ……polyethylene glycol
**PG-1** ……protegrin-1
**p-TEFb** ……positive transcription elongation factor b
**RBD** ……RNA-binding domain
**Rev** ……Regulator of Virion
**$R_f$** ……reference factor
**Rh** ……rhodamine
**RNA** ……ribonucleic acid
**RNAi** ……RNA interference
**RNP** ……ribonucleoprotein
**RP** ……reverse phase
**RRE** ……Rev response element
**RRM** ……RNA-recognition motif
**rRNA** ……ribosomal RNA
**RT** ……reverse transcriptase
**rt** ……room temperature

**SAR** ……..structure activity relationship
**siRNAs** ……small interfering RNAs
**SPPS** ……..…….solid phase peptide synthesis
**TAK** ……..Tat-activated kinase
**TAR** ……transactivator response element
**Tat** ……..……transcriptional transactivator
**TEMED** ……..……*N*,*N*,*N*′,*N*′-tetramethyl ethylene diamine
**TF** ……..transcription factor
**Tf** ……..trifluoromethane sulfonyl
**TFA** ……trifluoroacetic acid
**TIPS** ……..triisopropylsilane
**THF** ……...……tetrahydrofuran
**TLC** ……..……..thin layer chromatography
**tRNA** ……..transporter RNA
**TOCSY** ……..……..totally correlated spectra
**Tris** ……tris-(hydroxymethyl)-aminomethane
**Trt** ……trityl
**wt** ……wild type
**ZnF** ……zinc finger

# Appendix 2: Analytical data of the BIV/HIV Tat peptidomimetics

**Table 47.** Summary of the HPLC and MS analytical data of BIV Tat peptidomimetics. Amino acids 1 and 12 are attached to a D-Pro-L-Pro template (with an exception for wt BIV Tat). MWcalc.= theoretical MW calculated from molecular formula, MWobs.=MW from MS.

| Mimetic | Position | | | | | | | | | | | | MW calc. | m/z ESI-MS | MW obser. ES-MS | $r_t$ [min] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | *7* | **8** | *9* | **10** | *11* | **12** | | | | |
| **wt BIV Tat** | Ac-SGPRPRGTRGKGRRIRR-$NH_2$ | | | | | | | | | | | | 2004.3 | 669.2[M+3H]$^{3+}$<br>502.1[M+4H]$^{4+}$ | 2004.4 | 5.87* |
| **BIV-2** | R | V | R | T | R | G | K | R | R | I | R | V | 1729.1 | 577,6[M+3H]$^{3+}$<br>433,7[M+4H]$^{4+}$ | 1729,8 | 7,33* |
| **L1-01** | R | V | R | T | R | G | K | R | R | I | R | L | 1743.2 | 581.8[M+3H]$^{3+}$<br>436.7[M+4H]$^{4+}$ | 1742.4 | 4.57** |
| **L1-02** | R | V | R | T | R | G | K | R | R | I | R | I | 1743.2 | 581.8[M+3H]$^{3+}$<br>436.8[M+4H]$^{4+}$ | 1742.4 | 4.39** |
| **L1-03** | R | V | R | T | R | G | K | R | R | I | R | F | 1777.2 | 593.0[M+3H]$^{3+}$<br>445.0[M+4H]$^{4+}$ | 1776.0 | 4.68** |
| **L1-04** | R | V | R | T | R | G | K | R | R | I | R | T | 1731.1 | 577.8[M+3H]$^{3+}$<br>433.8[M+4H]$^{4+}$ | 1730.4 | 3.55** |
| **L1-05** | R | V | R | T | R | G | K | R | R | I | R | N | 1744.1 | 582.1[M+3H]$^{3+}$<br>437.0[M+4H]$^{4+}$ | 1743.3 | 3.32** |
| **L1-06** | R | V | R | T | R | G | K | R | R | I | R | R | 1786.2 | 596.0[M+3H]$^{3+}$<br>447.3[M+4H]$^{4+}$ | 1785.2 | 3.36** |
| **L1-07** | R | T | R | T | R | G | K | R | R | I | R | V | 1731.1 | 577.9[M+3H]$^{3+}$<br>433.7[M+4H]$^{4+}$ | 1730.7 | 4.04** |
| **L1-08** | R | N | R | T | R | G | K | R | R | I | R | V | 1744.1 | 582.1[M+3H]$^{3+}$<br>437.0[M+4H]$^{4+}$ | 1743.3 | 4.07** |
| **L1-09** | R | V | R | T | R | G | K | R | R | L | R | V | 1729.1 | 577.0[M+3H]$^{3+}$<br>433.2[M+4H]$^{4+}$ | 1728.0 | 4.28** |
| **L1-10** | R | V | R | T | R | G | K | R | R | X | R | V | 1769.0 | 590.5[M+3H]$^{3+}$<br>443.3[M+4H]$^{4+}$ | 1768.5 | 4.87** |
| **L1-11** | R | V | R | T | R | G | K | R | R | F | R | V | 1763.1 | 588.5[M+3H]$^{3+}$<br>441.9[M+4H]$^{4+}$ | 1762.5 | 4.36** |
| **L1-12** | R | V | R | T | R | G | K | R | R | Y | R | V | 1779.1 | 593.8[M+3H]$^{3+}$<br>445.8[M+4H]$^{4+}$ | 1778.4 | 3.98** |
| **L1-13** | R | V | R | T | R | G | K | R | R | N | R | V | 1730.1 | 577.4[M+3H]$^{3+}$<br>433.5[M+4H]$^{4+}$ | 1729.2 | 3.68** |
| **L1-14** | R | V | R | T | R | G | K | R | R | Q | R | V | 1744.1 | 582.4[M+3H]$^{3+}$<br>437.0[M+4H]$^{4+}$ | 1744.2 | 3.62** |
| **L1-15** | R | V | R | Q | R | G | K | R | R | I | R | V | 1756.1 | 586.4[M+3H]$^{3+}$<br>440.1[M+4H]$^{4+}$ | 1756.2 | 4.19** |
| **L1-16** | R | V | R | V | R | G | K | R | R | I | R | V | 1727.1 | 576.4[M+3H]$^{3+}$<br>432.6[M+4H]$^{4+}$ | 1726.2 | 4.17** |
| **L1-17** | R | V | R | Y | R | G | K | R | R | I | R | V | 1791.2 | 597.8[M+3H]$^{3+}$<br>448.7[M+4H]$^{4+}$ | 1790.4 | 4.07** |
| **L1-18** | R | V | R | K | R | G | K | R | R | I | R | V | 1756.2 | 586.4[M+3H]$^{3+}$<br>440.0[M+4H]$^{4+}$ | 1756.2 | 4.18** |
| **L1-19** | R | V | K | T | R | G | K | R | R | I | R | V | 1701.1 | 567.9[M+3H]$^{3+}$<br>426.4[M+4H]$^{4+}$ | 1700.7 | 4.04** |
| **L1-20** | R | V | U | T | R | G | K | R | R | I | R | V | 1730.1 | 577.5[M+3H]$^{3+}$<br>433.4[M+4H]$^{4+}$ | 1729.5 | 4.10** |
| **L1-21** | R | V | O | T | R | G | K | R | R | I | R | V | 1687.2 | 563.6[M+3H]$^{3+}$<br>422.7[M+4H]$^{4+}$ | 1686.8 | 5.79* |
| **L1-22** | R | V | R | T | K | G | K | R | R | I | R | V | 1701.1 | 567.9[M+3H]$^{3+}$<br>426.4[M+4H]$^{4+}$ | 1700.7 | 4.07** |
| **L1-23** | R | V | R | T | U | G | K | R | R | I | R | V | 1730.1 | 577.4[M+3H]$^{3+}$<br>433.5[M+4H]$^{4+}$ | 1729.2 | 4.08** |
| **L1-24** | R | V | R | T | O | G | K | R | R | I | R | V | 1687.1 | 563.2[M+3H]$^{3+}$<br>422.7[M+4H]$^{4+}$ | 1686.6 | 4.07** |
| **L1-25** | R | V | R | T | N | G | K | R | R | I | R | V | 1687.0 | 563.1[M+3H]$^{3+}$<br>422.8[M+4H]$^{4+}$ | 1686.3 | 4.09** |
| **L1-26** | R | V | R | T | Q | G | K | R | R | I | R | V | 1701.1 | 567.9[M+3H]$^{3+}$ | 1700.7 | 4.04** |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **L1-27** | R | V | R | T | I | G | K | R | R | I | R | V | 1686.1 | 426.3[M+4H]$^{4+}$<br>562.8[M+3H]$^{3+}$ | 1685.4 | 5.07** |
| **L1-28** | R | V | R | T | L | G | K | R | R | I | R | V | 1686.1 | 422.5[M+4H]$^{4+}$<br>526.9[M+3H]$^{3+}$ | 1685.7 | 5.13** |
| **L1-29** | R | V | R | T | Y | G | K | R | R | I | R | V | 1736.1 | 422.5[M+4H]$^{4+}$<br>579.4[M+3H]$^{3+}$ | 1735.2 | 4.72** |
| **L1-30** | R | V | R | T | R | G | K | K | R | I | R | V | 1701.1 | 435.0[M+4H]$^{4+}$<br>567.9[M+3H]$^{3+}$ | 1700.7 | 4.01** |
| **L1-31** | R | V | R | T | R | G | K | U | R | I | R | V | 1730.1 | 426.4[M+4H]$^{4+}$<br>577.5[M+3H]$^{3+}$ | 1729.5 | 4.16** |
| **L1-32** | R | V | R | T | R | G | K | O | R | I | R | V | 1687.1 | 433.5[M+4H]$^{4+}$<br>563.2[M+3H]$^{3+}$ | 1686.6 | 3.98** |
| **L1-33** | R | V | R | T | R | G | K | N | R | I | R | V | 1687.0 | 422.6[M+4H]$^{4+}$<br>563.1[M+3H]$^{3+}$ | 1686.3 | 4.26** |
| **L1-34** | R | V | R | T | R | G | K | F | R | I | R | V | 1720.1 | 422.7[M+4H]$^{4+}$<br>540.6[M+3H]$^{3+}$ | 1618.8 | 5.10** |
| **L1-35** | R | V | R | T | R | G | K | Y | R | I | R | V | 1736.1 | 405.7[M+4H]$^{4+}$<br>579.4[M+3H]$^{3+}$ | 1735.2 | 4.89** |
| **L1-36** | R | V | R | T | R | G | K | R | K | I | R | V | 1701.1 | 435.0[M+4H]$^{4+}$<br>567.9[M+3H]$^{3+}$ | 1700.7 | 4.12** |
| **L1-37** | R | V | R | T | R | G | K | R | U | I | R | V | 1730.1 | 426.4[M+4H]$^{4+}$<br>577.5[M+3H]$^{3+}$ | 1729.5 | 4.26** |
| **L1-38** | R | V | R | T | R | G | K | R | O | I | R | V | 1687.1 | 433.4[M+4H]$^{4+}$<br>563.1[M+3H]$^{3+}$ | 1686.3 | 4.36** |
| **L1-39** | R | V | R | T | R | G | K | R | N | I | R | V | 1687.0 | 422.7[M+4H]$^{4+}$<br>563.2[M+3H]$^{3+}$ | 1686.6 | 4.12** |
| **L1-40** | R | V | R | T | R | G | K | R | Q | I | R | V | 1701.1 | 422.8[M+4H]$^{4+}$<br>567.9[M+3H]$^{3+}$ | 1700.7 | 4.24** |
| **L1-41** | R | V | R | T | R | G | K | R | R | I | K | V | 1701.1 | 426.4[M+4H]$^{4+}$<br>567.9[M+3H]$^{3+}$ | 1700.7 | 4.02** |
| **L1-42** | R | V | R | T | R | G | K | R | R | I | U | V | 1730.1 | 426.5[M+4H]$^{4+}$<br>577.4[M+3H]$^{3+}$ | 1729.2 | 4.24** |
| **L1-43** | R | V | R | T | R | G | K | R | R | I | O | V | 1687.1 | 433.4[M+4H]$^{4+}$<br>563.1[M+3H]$^{3+}$ | 1686.3 | 3.99** |
| **L1-44** | R | V | R | T | R | G | K | R | R | I | N | V | 1687.0 | 422.7[M+4H]$^{4+}$<br>563.1[M+3H]$^{3+}$ | 1686.3 | 4.26** |
| **L1-45** | R | V | R | T | R | G | K | R | R | I | Q | V | 1701.1 | 422.7[M+4H]$^{4+}$<br>567.8[M+3H]$^{3+}$ | 1700.4 | 4.30** |
| | | | | | | | | | | | | | | 426.4[M+4H]$^{4+}$ | | |
| **L2-01** | R | V | R | T | R | G | K | R | R | I | R | I | 1743.2 | 581.8[M+3H]$^{3+}$<br>436.8[M+4H]$^{4+}$ | 1742.4 | 4.39** |
| **L2-02** | U | V | R | T | R | G | K | R | R | I | R | I | 1744.2 | 582.5[M+3H]$^{3+}$<br>437.2[M+4H]$^{4+}$ | 1774.5 | 7.84* |
| **L2-03** | K | V | R | T | R | G | K | R | R | I | R | I | 1715.1 | 573.1[M+3H]$^{3+}$<br>429.9[M+4H]$^{4+}$ | 1715.6 | 7.60* |
| **L2-04** | O | V | R | T | R | G | K | R | R | I | R | I | 1701.4 | 568.0[M+3H]$^{3+}$<br>426.6[M+4H]$^{4+}$ | 1701.0 | 7.38* |
| **L2-05** | N | V | R | T | R | G | K | R | R | I | R | I | 1701.1 | 568.3[M+3H]$^{3+}$<br>426.3[M+4H]$^{4+}$ | 1701.9 | 7.67* |
| **L2-06** | Q | V | R | T | R | G | K | R | R | I | R | I | 1715.1 | 572.7[M+3H]$^{3+}$<br>429.8[M+4H]$^{4+}$ | 1715.1 | 7.80* |
| **L2-07** | Y | V | R | T | R | G | K | R | R | I | R | I | 1750.1 | 584.2[M+3H]$^{3+}$<br>438.6[M+4H]$^{4+}$ | 1749.6 | 9.03* |
| **L2-08** | L | V | R | T | R | G | K | R | R | I | R | I | 1700.1 | 567.6[M+3H]$^{3+}$<br>426.1[M+4H]$^{4+}$ | 1699.8 | 9.38* |
| **L2-09** | W | V | R | T | R | G | K | R | R | I | R | I | 1773.2 | 592.2[M+3H]$^{3+}$<br>444.4[M+4H]$^{4+}$ | 1773.6 | 9.98* |
| **L2-10** | R | Q | R | T | R | G | K | R | R | I | R | I | 1772.2 | 591.9[M+3H]$^{3+}$<br>444.2[M+4H]$^{4+}$ | 1772.7 | 7.60* |
| **L2-11** | R | Y | R | T | R | G | K | R | R | I | R | I | 1807.2 | 603.4[M+3H]$^{3+}$<br>453.0[M+4H]$^{4+}$ | 1807.2 | 7.42* |
| **L2-12** | R | W | R | T | R | G | K | R | R | I | R | I | 1830.2 | 611.2[M+3H]$^{3+}$<br>458.7[M+4H]$^{4+}$ | 1830.6 | 7.49* |
| **L2-13** | R | L | R | T | R | G | K | R | R | I | R | I | 1757.2 | 587.2[M+3H]$^{3+}$<br>440.5[M+4H]$^{4+}$ | 1758.0 | 7.73* |
| **L2-14** | R | V | R | T | R | N | G | R | R | I | R | I | 1729.1 | 577.6[M+3H]$^{3+}$<br>433.5[M+4H]$^{4+}$ | 1729.8 | 7.04* |
| **L2-15** | R | V | R | T | R | K | G | R | R | I | R | I | 1743.2 | 582.2[M+3H]$^{3+}$ | 1743.6 | 7.24* |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **L2-16** | R | V | R | T | R | G | R | R | R | I | R | I | 1771.2 | 436.8[M+4H]$^{4+}$<br>591.3[M+3H]$^{3+}$ | 1770.9 | 7.20* |
| **L2-17** | R | V | R | T | R | G | Y | R | R | I | R | I | 1778.2 | 444.0[M+4H]$^{4+}$<br>593.8[M+3H]$^{3+}$ | 1778.4 | 7.65* |
| **L2-18** | R | V | R | T | R | G | Q | R | R | I | R | I | 1743.1 | 445.7[M+4H]$^{4+}$<br>582.1[M+3H]$^{3+}$ | 1743.3 | 7.36* |
| **L2-19** | R | V | R | T | R | G | N | R | R | I | R | I | 1729.1 | 436.9[M+4H]$^{4+}$<br>577.5[M+3H]$^{3+}$ | 1729.5 | 7.26* |
| **L2-20** | R | V | R | T | R | G | O | R | R | I | R | I | 1729.1 | 433.4[M+4H]$^{4+}$<br>577.3[M+3H]$^{3+}$ | 1728.9 | 7.55* |
| **L2-21** | R | V | R | T | R | G | U | R | R | I | R | I | 1772.2 | 433.4[M+4H]$^{4+}$<br>591.8[M+3H]$^{3+}$ | 1772.4 | 7.99* |
| **L2-22** | R | V | Q | T | R | G | K | R | R | I | R | I | 1715.1 | 444.2[M+4H]$^{4+}$<br>572.5[M+3H]$^{3+}$ | 1714.5 | 7.60* |
| **L2-23** | R | V | N | T | R | G | K | R | R | I | R | I | 1701.1 | 429.9[M+4H]$^{4+}$<br>567.8[M+3H]$^{3+}$ | 1700.4 | 7.81* |
| **L2-24** | R | T | R | T | R | G | K | R | R | I | R | I | 1745.1 | 426.3[M+4H]$^{4+}$<br>582.8[M+3H]$^{3+}$ | 1745.4 | 7.84* |
| **L2-25** | R | N | R | T | R | G | K | R | R | I | R | I | 1758.1 | 437.6[M+4H]$^{4+}$<br>587.1[M+3H]$^{3+}$ | 1758.3 | 7.81* |
| **L2-26** | R | V | R | T | R | G | K | R | R | I | O | I | 1701.1 | 440.7[M+4H]$^{4+}$<br>568.1[M+3H]$^{3+}$ | 1701.3 | 7.57* |
| **L2-27** | R | V | R | T | R | G | K | R | R | I | K | I | 1715.1 | 426.4[M+4H]$^{4+}$<br>572.8[M+3H]$^{3+}$ | 1715.4 | 7.54* |
| **L2-28** | R | V | R | T | R | G | K | R | K | I | O | I | 1673.1 | 430.0[M+4H]$^{4+}$<br>559.2[M+3H]$^{3+}$ | 1673.6 | 7.33* |
| **L2-29** | R | V | R | Q | R | G | K | R | R | I | R | I | 1770.2 | 419.4[M+4H]$^{4+}$<br>591.1[M+3H]$^{3+}$ | 1770.3 | 7.53* |
| **L2-30** | R | V | R | V | R | G | K | R | R | I | R | I | 1741.2 | 443.6[M+4H]$^{4+}$<br>581.4[M+3H]$^{3+}$ | 1741.2 | 7.49* |
| **L2-31** | R | V | R | Y | R | G | K | R | R | I | R | I | 1805.2 | 436.4[M+4H]$^{4+}$<br>603.0[M+3H]$^{3+}$ | 1806.0 | 7.41* |
| **L2-32** | R | T | R | Q | R | G | K | R | R | I | R | I | 1772.2 | 452.3[M+4H]$^{4+}$<br>591.6[M+3H]$^{3+}$ | 1771.8 | 7.53* |
| **L2-33** | R | T | R | T | R | G | K | R | R | I | O | I | 1703.1 | 444.1[M+4H]$^{4+}$<br>568.5[M+3H]$^{3+}$ | 1702.5 | 7.73* |
| **L2-34** | R | T | R | T | R | G | K | R | K | I | O | I | 1675.1 | 426.9[M+4H]$^{4+}$<br>559.1[M+3H]$^{3+}$ | 1674.3 | 7.31* |
| **L2-35** | R | T | R | Q | R | G | K | R | R | I | O | I | 1730.1 | 419.9[M+4H]$^{4+}$<br>577.9[M+3H]$^{3+}$ | 1730.7 | 7.30* |
| **L2-36** | R | T | R | V | R | G | K | R | R | I | O | I | 1701.1 | 433.7[M+4H]$^{4+}$<br>568.5[M+3H]$^{3+}$ | 1701.6 | 7.35* |
| **L2-37** | R | T | R | Y | R | G | K | R | R | I | O | I | 1765.1 | 426.4[M+4H]$^{4+}$<br>589.8[M+3H]$^{3+}$ | 1764.4 | 7.30* |
| **L2-38** | R | V | R | T | R | G | K | R | R | F | R | I | 1777.2 | 442.1[M+4H]$^{4+}$<br>593.4[M+3H]$^{3+}$ | 1777.2 | 7.64* |
| **L2-39** | R | T | R | T | R | G | K | R | R | F | R | I | 1779.1 | 445.4[M+4H]$^{4+}$<br>594.7[M+3H]$^{3+}$ | 1800.0 | 7.78* |
| **L2-40** | R | T | R | T | R | G | K | R | R | F | O | I | 1737.1 | 446.0[M+4H]$^{4+}$<br>580.2[M+3H]$^{3+}$ | 1737.6 | 7.30* |
| **L2-41** | R | V | R | Q | R | G | K | R | R | F | O | I | 1762.2 | 435.6[M+4H]$^{4+}$<br>588.4[M+3H]$^{3+}$ | 1762.2 | 7.58* |
| **L2-42** | R | T | R | V | R | G | K | R | K | F | O | I | 1707.1 | 441.6[M+4H]$^{4+}$<br>570.0[M+3H]$^{3+}$ | 1707.0 | 7.55* |
| **L2-43** | R | T | R | Y | R | G | K | R | K | F | O | I | 1771.1 | 428.1[M+4H]$^{4+}$<br>591.5[M+3H]$^{3+}$ | 1771.5 | 7.50* |
| | | | | | | | | | | | | | | 444.0[M+4H]$^{4+}$ | | |

* analytical column: Agilent, Zorbax Eclipse XDB-C18 4.6 x 250 mm (5 µm), 1 ml/min, 5CV, 10-90% MeCN in $H_2O$ +0,1%TFA; ** analytical column: Phenomenex, Jupiter Proteo 90A, 2.0 x 50 mm(4 µm), 0,5 ml/min, 8 min, 5-60% MeCN in $H_2O$ +0,1%TFA

**Table 48.** Summary of the HPLC and MS analytical data of HIV Tat peptidomimetics. Amino acids 1 and 12 are attached to a D-Pro-L-Pro template (with an exception for wt HIV Tat). MWcalc.= theoretical MW calculated from molecular formula, MWobs.=MW from MS.

| Mimetic | Position | | | | | | | | | | | | Template | MW calc | m/z ESI-MS | $MW_{obs}$ ES-MS | $r_t$ [min] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | *7* | **8** | *9* | **10** | *11* | **12** | *13*-**14** | | | | |
| **wt HIV-1 Tat** | Ac-SFTTKALGISYGRKKRRQRRRPPQGSQTHQVSLSKQ-$NH_2$ | | | | | | | | | | | | | 4210.9 | 1404.4[M+3H]$^{3+}$ 1054.0[M+4H]$^{4+}$ | 4210.2 | 8.00 |
| **L3-01** | R | V | R | T | R | K | G | R | R | I | R | I | $^{D}$R-$^{L}$P | 1802.2 | 601.5[M+3H]$^{3+}$ 451.5[M+4H]$^{4+}$ | 1801.5 | 6.40 |
| **L3-02** | R | V | R | T | R | K | G | R | R | R | R | I | $^{D}$P-$^{L}$P | 1786.2 | 893.7[M+2H]$^{2+}$ 596.5[M+3H]$^{3+}$ | 1786.5 | 7.06 |
| **L3-03** | R | C | R | T | R | K | G | R | R | R | C | I | $^{D}$P-$^{L}$P | 1735.2 | 868.6[M+2H]$^{2+}$ 579.4[M+3H]$^{3+}$ | 1734.4 | 7.43 |
| **L3-04** | R | V | R | T | R | R | G | R | R | I | R | I | $^{D}$P-$^{L}$P | 1771.2 | 886.4[M+2H]$^{2+}$ 592.0[M+3H]$^{3+}$ | 1770.8 | 7.67 |
| **L3-05** | R | C | R | T | R | K | G | R | R | K | C | I | $^{D}$P-$^{L}$P | 1707.1 | 854.3[M+2H]$^{2+}$ 569.7[M+3H]$^{4+}$ | 1706.6 | 7.22 |
| **L3-06** | R | V | R | T | R | hR | G | R | R | K | R | I | $^{D}$P-$^{L}$P | 1800.2 | 900.8[M+2H]$^{2+}$ 600.8[M+3H]$^{3+}$ | 1799.6 | 7.10 |

Analytical column: Agilent, Zorbax Eclipse XDB-C18 4.6 x250 mm (5 µm), 5CV, 1 ml/min; 10-90% MeCN in $H_2O$ +0,1%TFA

**Table 49** The HPLC and MS analytical data of L2-15a.

| Mimetic | Position | | | | | | | | | | | | MW calc | m/z MALDI | $r_t$ [min] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | *7* | **8** | *9* | **10** | *11* | **12** | | | |
| **L2-15a** | Ac-R | V | R | T | R | K | G | R | R | I | R | I-$NH_2$ | 1608 | 1608.0 | 6.94 |

Analytical column: Agilent, Zorbax Eclipse XDB-C18 4.6 x250 mm (5 µm), 5CV, 1 ml/min; 10-90% MeCN in $H_2O$ +0,1%TFA

**Table 50.** Summary of the HPLC and MS analytical data of L2-15hR peptidomimetics [Single mutations of Arginine to homoArginine (hR) in L2-15 peptide]. Amino acids 1 and 12 are attached to a D-Pro-L-Pro template. MWcalc.= theoretical MW calculated from molecular formula, MWobs.=MW from MS.

| Mimetic | Position | | | | | | | | | | | | MW calc | m/z ESI-MS | $MW_{obs}$ ESI-MS | $r_t$ [min] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | *7* | **8** | *9* | **10** | *11* | **12** | | | | |
| **L2-15hR01** | hR | V | R | T | R | K | G | R | R | I | R | I | 1757.2 | 587.1[M+3H]$^{3+}$ 440.4[M+4H]$^{4+}$ | 1757.6 | 15.46 |
| **L2-15hR02** | R | V | hR | T | R | K | G | R | R | I | R | I | 1757.2 | 587.0[M+3H]$^{3+}$ 440.4[M+4H]$^{4+}$ | 1758.0 | 15.29 |
| **L2-15hR03** | R | V | R | T | hR | K | G | R | R | I | R | I | 1757.2 | 587.1[M+3H]$^{3+}$ 440.8[M+4H]$^{3+}$ | 1758.3 | 15.01 |
| **L2-15hR04** | R | V | R | T | R | K | G | hR | R | I | R | I | 1757.2 | 587.1[M+3H]$^{3+}$ 440.8[M+4H]$^{4+}$ | 1758.3 | 15.39 |
| **L2-15hR05** | R | V | R | T | R | K | G | R | hR | I | R | I | 1757.2 | 586.4[M+3H]$^{3+}$ 440.4[M+4H]$^{4+}$ | 1756.2 | 15.22 |
| **L2-15hR06** | R | V | R | T | R | K | G | R | R | I | hR | I | 1757.2 | 586.4[M+3H]$^{3+}$ 440.4[M+4H]$^{4+}$ | 1756.2 | 15.38 |

Analytical column: Agilent, Zorbax Eclipse XDB-C18 4.6 x 250 mm (5 µm), 6CV, 1 ml/min; 10-50% MeCN in $H_2O$ +0,1%TFA

**Table 51.** Summary of the HPLC and MS analytical data of HIV Tat peptidomimetics [L4-Library]. Amino acids 1 and 16 (13/14) are attached to a D-Pro-L-Pro template. MWcalc.= theoretical MW calculated from molecular formula, MWobs.=MW from MS.

| **Mimetic** | **Position** | | | | | | | | | | | | | | | | **MW calc** | **m/z ES-MS** | **$MW_{obs}$ ESI-MS** | **$r_t$ [min]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | **7** | *8* | **9** | *10* | **11** | *12* | **13** | *14* | **15** | *16* | | | | |
| **L4-01** | R | V | R | C | R | T | R | G | K | R | R | R | C | A | R | V | 2203.7 | 552.1$[M+4H]^{4+}$<br>442.0$[M+5H]^{5+}$ | 2204.4 | 7.20 |
| **L4-02** | R | V | R | C | R | Q | R | K | G | R | A | R | C | I | R | I | 2201.7 | 551.5$[M+4H]^{4+}$<br>441.5$[M+5H]^{5+}$ | 2202.0 | 8.52 |
| **L4-03** | R | V | R | C | R | K | L | K | G | Q | T | R | C | I | R | I | 2160.7 | 721.4$[M+3H]^{3+}$<br>541.5$[M+4H]^{4+}$ | 2161.2 | 8.32 |
| **L4-04** | R | V | R | C | R | K | R | $^{D}$R | G | Q | T | R | C | I | R | I | 2231.8 | 744.7$[M+3H]^{3+}$<br>559.3$[M+4H]^{4+}$ | 2231.1 | 8.63 |
| **L4-05** | R | V | R | C | R | Q | R | G | P | G | K | R | C | I | R | I | 2128.6 | 710.1$[M+3H]^{3+}$<br>533.2$[M+4H]^{4+}$ | 2128.8 | 7.93 |
| **L4-06** | R | V | R | C | R | K | Q | G | P | G | T | R | C | I | R | I | 2073.6 | 692.4$[M+3H]^{3+}$<br>519.4$[M+4H]^{4+}$ | 2074.2 | 7.91 |
| **L4-07** | R | V | R | C | R | R | K | G | P | G | Q | R | C | I | R | I | 2128.6 | 711.2$[M+3H]^{3+}$<br>533.4$[M+4H]^{4+}$ | 2129.6 | 7.48 |
| **L4-08** | R | V | R | C | R | K | R | G | P | G | Q | R | C | I | R | I | 2128.6 | 710.7$[M+3H]^{3+}$<br>533.5$[M+4H]^{4+}$ | 2129.1 | 7.93 |
| **L4-09** | R | C | R | V | R | Q | R | K | G | R | A | R | K | I | C | I | 2173.7 | 726.8$[M+3H]^{3+}$<br>544.8$[M+4H]^{4+}$ | 2175.2 | 7.99 |
| **L4-10** | R | C | R | V | R | K | Q | G | P | G | T | R | K | I | C | I | 2045.5 | 683.3$[M+3H]^{3+}$<br>512.3$[M+4H]^{4+}$ | 2045.2 | 7.61 |
| **L4-11** | R | C | R | V | R | K | Q | G | P | G | K | R | K | I | C | I | 2072.6 | 692.5$[M+3H]^{3+}$<br>519.6$[M+4H]^{4+}$ | 2074.4 | 7.96 |
| **L4-12** | R | C | R | V | R | K | R | G | P | G | Q | R | K | I | C | I | 2100.6 | 701.7$[M+3H]^{3+}$<br>526.5$[M+4H]^{4+}$ | 2102.0 | 7.70 |
| **L4-13** | R | C | R | V | R | R | K | G | P | G | A | R | K | I | C | I | 2043.6 | 682.5$[M+3H]^{3+}$<br>511.9$[M+4H]^{4+}$ | 2044.5 | 8.44 |
| **L4-14** | R | V | R | C | R | K | K | G | K | C | I | R | I | | | | 1790.3 | 597.8$[M+3H]^{3+}$<br>449.4$[M+4H]^{4+}$ | 1790.4 | 7.84 |
| **L4-15** | R | V | R | C | R | R | R | G | R | C | I | R | I | | | | 1874.4 | 625.9$[M+3H]^{3+}$<br>470.1$[M+4H]^{4+}$ | 1874.7 | 7.90 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **L4-16** | R | V | R | T | R | K | K | G | K | R | I | R | I | | 1843.3 | 615.7[M+3H]$^{3+}$ 462.2[M+4H]$^{4+}$ | 1844.1 | 7.62 |
| **L4-17** | R | V | R | C | R | K | K | $^{D}$P | K | C | I | R | I | | 1830.4 | 611.5[M+3H]$^{3+}$ 459.1[M+4H]$^{4+}$ | 1831.5 | 7.82 |
| **L4-18** | R | V | R | T | R | K | K | $^{D}$P | K | R | I | R | I | | 1883.4 | 629.0[M+3H]$^{3+}$ 472.2[M+4H]$^{4+}$ | 1884.0 | 7.54 |
| **L4-19** | R | V | R | V | R | K | K | G | R | T | I | R | I | | 1814.3 | 606.0[M+3H]$^{3+}$ 454.9[M+4H]$^{4+}$ | 1815.0 | 8.37 |
| **L4-20** | R | V | R | C | R | K | R | K | G | R | C | I | R | I | 1974.5 | 659.6[M+3H]$^{3+}$ 494.9[M+4H]$^{4+}$ | 1975.6 | 8.25 |
| **L4-21** | R | V | R | C | R | R | K | R | G | R | C | I | R | I | 2002.5 | 668.4[M+3H]$^{3+}$ 501.7[M+4H]$^{4+}$ | 2002.2 | 8.29 |
| **L4-22** | R | V | R | T | R | Q | R | K | G | R | R | I | R | I | 2027.5 | 676.9[M+3H]$^{3+}$ 508.1[M+4H]$^{4+}$ | 2027.7 | 8.10 |
| **L4-23** | R | V | R | V | R | K | R | K | G | R | Q | I | R | I | 1997.5 | 666.9[M+3H]$^{3+}$ 500.7[M+4H]$^{4+}$ | 1997.7 | 8.25 |
| **L4-24** | R | V | R | Q | R | K | R | K | G | R | V | I | R | I | 1997.5 | 666.9[M+3H]$^{3+}$ 500.6[M+4H]$^{4+}$ | 1997.7 | 8.02 |
| **L4-25** | R | V | R | C | R | K | R | K | $^{D}$P | R | C | I | R | I | 2014.6 | 673.5[M+3H]$^{3+}$ 504.8[M+4H]$^{4+}$ | 2015.2 | 8.25 |

Analytical column: Agilent, Zorbax Eclipse XDB-C18 4.6 x 250 mm (5 μm), 6CV, 1 ml/min 10-50% MeCN in $H_2O$ +0,1%TFA

**Table 52.** Summary of the HPLC and MS analytical data of HIV Tat peptidomimetics [L5-Library]. Amino acids 1 and 16 are attached to a D-Pro-L-Pro template. MWcalc.= theoretical MW calculated from molecular formula, MWobs.=MW from MS.

| **Mimetic** | **Position** | | | | | | | | | | | | | | | | **MW calc** | **m/z ESI-MS** | **$MW_{obs}$ ESI-MS** | **$r_t$ [min]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | **7** | *8* | **9** | *10* | **11** | *12* | **13** | *14* | **15** | *16* | | | | |
| **L5-01** | R | V | R | C | R | K | R | R | G | Q | T | R | C | I | R | I | 2233.8 | 745.3$[M+3H]^{3+}$<br>559.0$[M+4H]^{4+}$ | 2232.9 | 15.10 |
| **L5-02** | R | V | K | C | R | K | R | K | G | Q | T | R | C | I | R | I | 2203.8 | 735.9$[M+3H]^{3+}$<br>552.3$[M+4H]^{4+}$ | 2204.7 | 15.02 |
| **L5-03** | R | V | R | C | R | K | L | $^{D}$R | G | Q | T | R | C | I | R | I | 2188.7 | 730.9$[M+3H]^{3+}$<br>548.5$[M+4H]^{4+}$ | 2189.7 | 16.42 |
| **L5-04** | R | V | R | C | R | R | K | $^{D}$R | G | Q | T | R | C | I | R | I | 2233.8 | 745.4$[M+3H]^{3+}$<br>559.2$[M+4H]^{4+}$ | 2233.2 | 14.93 |
| **L5-05** | R | V | R | T | R | K | R | $^{D}$R | G | Q | T | R | A | I | K | I | 2171.6 | 725.4$[M+3H]^{3+}$<br>544.3$[M+4H]^{4+}$ | 2173.2 | 14.66 |
| **L5-06** | R | C | R | V | R | K | R | $^{D}$R | G | Q | T | R | R | I | C | I | 2231.8 | 744.3$[M+3H]^{3+}$<br>558.6$[M+4H]^{4+}$ | 2229.9 | 15.62 |
| **L5-07** | R | V | R | C | R | K | R | G | P | G | A | R | C | I | K | I | 2043.6 | 682.5$[M+3H]^{3+}$<br>512.1$[M+4H]^{4+}$ | 2044.5 | 15.56 |
| **L5-08** | R | C | R | V | R | R | K | K | G | R | A | R | K | I | C | I | 2175.8 | 725.9$[M+3H]^{3+}$<br>544.8$[M+4H]^{4+}$ | 2174.7 | 15.59 |
| **L5-09** | R | C | R | V | R | Q | R | G | P | G | A | R | K | I | C | I | 2043.5 | 682.8$[M+3H]^{3+}$<br>512.3$[M+4H]^{4+}$ | 2045.2 | 16.32 |
| **L5-10** | R | V | R | C | R | T | R | K | G | R | A | R | C | I | K | I | 2146.7 | 717.1$[M+3H]^{3+}$<br>538.0$[M+4H]^{4+}$ | 2148.0 | 15.39 |
| **L5-11** | R | V | R | C | R | Q | R | K | G | R | R | A | C | I | R | I | 2201.7 | 735.2$[M+3H]^{3+}$<br>551.5$[M+4H]^{4+}$ | 2202.6 | 15.35 |
| **L5-12** | R | V | R | C | R | Q | R | K | G | R | R | I | C | I | R | I | 2243.8 | 748.7$[M+3H]^{3+}$<br>561.6$[M+4H]^{4+}$ | 2243.1 | 17.75 |

Analytical column: Agilent, Zorbax Eclipse XDB-C18, 4.6 x250 mm (5 μm), 6CV, 1 ml/min 10-50% MeCN in $H_2O$ +0,1%TFA

**Table 53.** Summary of the HPLC and MS analytical data of HIV Tat peptidomimetics [Single mutations of arginine to homoarginine (hR) or arginine to lysine in L4-04]. Amino acids 1 and 16 are attached to a D-Pro-L-Pro template. MWcalc.= theoretical MW calculated from molecular formula, MWobs.=MW from MS.

| **Mimetic** | **Position** | | | | | | | | | | | | | | | | **MW calc** | **m/z ES-MS** | **$MW_{obs}$ ES-MS** | **$r_t$ [min]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | **7** | *8* | **9** | *10* | **11** | *12* | **13** | *14* | **15** | *16* | | | | |
| **L4-04hR01** | hR | V | R | C | R | K | R | $^{D}$R | G | Q | T | R | C | I | R | I | 2245.8 | 749.9[M+3H]$^{3+}$ 562.8[M+4H]$^{4}$ | 2246.7 | 16.76 |
| **L4-04hR02** | R | V | hR | C | R | K | R | $^{D}$R | G | Q | T | R | C | I | R | I | 2245.8 | 749.9[M+3H]$^{3+}$ 562.7[M+4H]$^{4}$ | 2246.7 | 16.53 |
| **L4-04hR03** | R | V | R | C | hR | K | R | $^{D}$R | G | Q | T | R | C | I | R | I | 2245.8 | 749.9[M+3H]$^{3+}$ 562.7[M+4H]$^{4}$ | 2246.7 | 16.49 |
| **L4-04hR04** | R | V | R | C | R | K | hR | $^{D}$R | G | Q | T | R | C | I | R | I | 2245.8 | 749.6[M+3H]$^{3+}$ 562.8[M+4H]$^{4+}$ | 2245.8 | 16.51 |
| **L4-04hR05** | R | V | R | C | R | K | R | $^{D}$R | G | Q | T | hR | C | I | R | I | 2245.8 | 750.4[M+3H]$^{3+}$ 562.7[M+4H]$^{4}$ | 2246.2 | 16.57 |
| **L4-04hR06** | R | V | R | C | R | K | R | $^{D}$R | G | Q | T | R | C | I | hR | I | 2245.8 | 750.3[M+3H]$^{3+}$ 562.7[M+4H]$^{4}$ | 2246.2 | 16.63 |
| **L4-04K01** | K | V | R | C | R | K | R | $^{D}$R | G | Q | T | R | C | I | R | I | 2203.8 | 736.1[M+3H]$^{3+}$ 552.0[M+4H]$^{4+}$ | 2204.4 | 17.52 |
| **L4-04K02** | R | V | K | C | R | K | R | $^{D}$R | G | Q | T | R | C | I | R | I | 2203.8 | 735.9[M+3H]$^{3+}$ 551.9[M+4H]$^{4+}$ | 2202.0 | 17.46 |
| **L4-04K03** | R | V | R | C | K | K | R | $^{D}$R | G | Q | T | R | C | I | R | I | 2203.8 | 735.9[M+3H]$^{3+}$ 551.7[M+4H]$^{4+}$ | 2161.2 | 17.44 |
| **L4-04K04** | R | V | R | C | R | K | K | $^{D}$R | G | Q | T | R | C | I | R | I | 2203.8 | 735.9[M+3H]$^{3+}$ 552.0[M+4H]$^{4+}$ | 2204.4 | 17.34 |
| **L4-04K05** | R | V | R | C | R | K | R | $^{D}$R | G | Q | T | K | C | I | R | I | 2203.8 | 736.0[M+3H]$^{3+}$ 552.3[M+4H]$^{4+}$ | 2204.4 | 17.27 |
| **L4-04K06** | R | V | R | C | R | K | R | $^{D}$R | G | Q | T | R | C | I | K | I | 2203.8 | 736.2[M+3H]$^{3+}$ 552.0[M+4H]$^{4+}$ | 2202.0 | 16.92 |
| **L4-04K07** | R | V | R | C | R | K | R | $^{D}$K | G | Q | T | R | C | I | R | I | 2203.8 | 735.8[M+3H]$^{3+}$ 551.8[M+4H]$^{4+}$ | 2204.4 | 15.74 |

Analytical column: Agilent, Zorbax Eclipse XDB-C18, 4.6 x250 mm (5 μm), 6CV, 1 ml/min 10-50% MeCN in $H_2O$ +0,1%TFA

**Table 54.** Summary of the HPLC and MS analytical data of HIV Tat peptidomimetics [Derivatives of L4-04]. MWcalc.= theoretical MW calculated from molecular formula, MWobs.=MW from MS. Abu=aminobutyric acid.

| **Mimetic** | **Position** | | | | | | | | | | | | | | | | | **MW calc** | **m/z ESI-MS** | **$MW_{obs}$ ESI-MS** | **$r_t$ [min]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | **7** | *8* | **9** | *10* | **11** | *12* | **13** | *14* | **15** | *16* | **17**-*18* | | | | |
| **L4-04a** | Ac-R | V | R | C | R | K | R | $^{D}$R | G | Q | T | R | C | I | R | I-$NH_2$ | - | 2096.6 | 700.5$[M+3H]^{3+}$<br>525.3$[M+4H]^{4+}$ | 2097.2 | 13.68 |
| **L4-04b** | R | V | R | Abu | R | K | R | $^{D}$R | G | Q | T | R | Abu | I | R | I | $^{D}$P-$^{L}$P | 2197.6 | 733.3$[M+3H]^{3+}$<br>550.8$[M+4H]^{4+}$ | 2196.9 | 16.01 |
| **L4-04c** | Ac-R | V | R | Abu | R | K | R | $^{D}$R | G | Q | T | R | Abu | I | R | I-$NH_2$ | - | 2062.5 | 688.8$[M+3H]^{3+}$<br>517.0$[M+4H]^{4+}$ | 2063.4 | 14.97 |

Analytical column: Agilent, Zorbax Eclipse XDB-C18, 4.6 x 250 mm (5 µm), 6CV, 1 ml/min 10-50% MeCN in $H_2O$ +0,1%TFA

# Appendix 3: Analytical data of the Rev peptidomimetics

**Table 55.** Summary of the HPLC and MS analytical data of Rev peptidomimetics. Amino acids 1 and 12 are attached to a D-Pro-L-Pro template. MWcalc.= theoretical MW calculated from molecular formula, MWobs.=MW from MS.

| Mimetic | Position | | | | | | | | | | | | MW calc | m/z ESI-MS | $MW_{obs}$ ESI-MS | $r_t$ [min] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | *7* | **8** | *9* | **10** | *11* | **12** | | | | |
| **BIV-5** | R | R | G | T | R | G | K | R | R | I | G | R | 1645.0 | 549.4[M+3H]$^{3+}$<br>412.3[M+4H]$^{4+}$ | 1645.2 | 6.56 |
| **R1-01** | W | R | R | R | A | T | $^{D}$R | Q | R | N | R | R | 1888.2 | 630.4[M+3H]$^{3+}$<br>472.9[M+4H]$^{4+}$ | 1888.2 | 7.45 |
| **R1-02** | W | R | R | R | A | P | $^{D}$R | Q | R | N | R | R | 1884.2 | 628.8[M+3H]$^{3+}$<br>471.9[M+4H]$^{4+}$ | 1883.4 | 7.52 |
| **R1-03** | W | R | R | R | G | P | $^{D}$R | Q | R | N | R | R | 1870.2 | 624.3[M+3H]$^{3+}$<br>468.5[M+4H]$^{4+}$ | 1869.9 | 7.48 |
| **R1-04** | W | R | R | R | V | P | $^{D}$R | Q | R | N | R | R | 1912.3 | 638.5[M+3H]$^{3+}$<br>479.0[M+4H]$^{4+}$ | 1912.5 | 7.69 |
| **R1-05** | W | R | R | R | A | G | $^{D}$R | Q | R | N | R | R | 1844.1 | 615.8[M+3H]$^{3+}$<br>461.9[M+4H]$^{4+}$ | 1844.4 | 7.40 |
| **R1-06** | W | R | R | R | A | G | K | Q | R | N | R | R | 1816.1 | 606.2[M+3H]$^{3+}$<br>455.0[M+4H]$^{4+}$ | 1815.6 | 7.28 |
| **R1-07** | W | R | R | R | A | T | $^{D}$R | G | R | N | R | R | 1817.1 | 606.3[M+3H]$^{3+}$<br>455.3[M+4H]$^{4+}$ | 1815.9 | 7.43 |
| **R1-08** | W | R | R | R | A | P | $^{D}$R | G | R | N | R | R | 1813.1 | 605.3[M+3H]$^{3+}$<br>454.3[M+4H]$^{4+}$ | 1812.9 | 7.51 |
| **R1-09** | W | R | R | R | G | P | $^{D}$R | G | R | N | R | R | 1799.1 | 600.7[M+3H]$^{3+}$<br>450.8[M+4H]$^{4+}$ | 1799.1 | 7.10 |
| **R1-10** | W | R | R | R | G | T | $^{D}$R | Q | R | N | R | R | 1874.2 | 625.7[M+3H]$^{3+}$<br>469.5[M+4H]$^{4+}$ | 1874.1 | 6.94 |
| **R1-11** | W | R | R | R | V | G | $^{D}$R | Q | R | N | R | R | 1872.2 | 625.2[M+3H]$^{3+}$<br>469.0[M+4H]$^{4+}$ | 1872.0 | 7.04 |
| **R1-12** | W | R | R | R | A | S | $^{D}$R | Q | R | N | R | R | 1874.2 | 625.4[M+3H]$^{3+}$<br>469.5[M+4H]$^{4+}$ | 1874.0 | 6.88 |
| **R2-01** | W | R | R | R | A | K | G | Q | R | N | R | R | 1816.1 | #1816.7[M+H]$^{+}$ | 1815.7 | 7.32 |
| **R2-02** | W | R | L | R | A | K | G | Q | R | N | R | R | 1773.1 | 887.3[M+2H]$^{2+}$ | 1772.6 | 8.07 |
| **R2-03** | W | R | Q | R | A | K | G | Q | R | N | R | R | 1788.1 | 894.7[M+2H]$^{2+}$<br>597.2[M+3H]$^{3+}$ | 1787.4 | 6.94 |
| **R2-04** | W | R | R | R | A | K | G | Q | R | N | R | V | 1759.1 | 880.3[M+2H]$^{2+}$ | 1758.6 | 7.12 |
| **R2-05** | W | R | L | R | A | $^{D}$R | K | G | R | N | R | R | 1801.2 | 901.5[M+2H]$^{2+}$ | 1801.0 | 7.84 |
| **R2-06** | W | R | L | R | A | K | K | G | R | N | R | R | 1773.1 | 887.8[M+2H]$^{2+}$ | 1773.6 | 7.96 |
| **R2-07** | W | R | K | R | A | G | K | Q | R | N | R | R | 1788.1 | 894.8[M+2H]$^{2+}$ | 1787.6 | 7.00 |
| **R2-08** | W | R | K | R | A | K | G | Q | R | N | R | R | 1788.1 | 894.8[M+2H]$^{2+}$ | 1787.6 | 7.10 |
| **R2-09** | W | R | R | R | A | $^{D}$R | G | Q | R | N | R | R | 1844.1 | 922.8[M+2H]$^{2+}$ | 1843.6 | 7.32 |
| **R2-10** | W | R | L | R | A | $^{D}$R | G | K | R | N | R | R | 1801.2 | 901.2[M+2H]$^{2+}$ | 1800.4 | 8.12 |
| **R2-11** | R | R | L | R | A | $^{D}$R | G | Q | R | N | R | R | 1771.1 | 886.6[M+2H]$^{2+}$ | 1771.2 | 7.41 |
| **R2-12** | R | R | L | R | A | K | G | Q | R | N | R | R | 1743.1 | 871.9[M+2H]$^{2+}$ | 1741.8 | 6.45 |
| **R2-13** | R | V | R | R | A | K | G | Q | R | N | R | R | 1729.1 | 865.5[M+2H]$^{2+}$ | 1729.0 | 5.98 |
| **R2-14** | R | C | R | R | A | K | G | Q | R | R | C | R | 1720.1 | #1720.7[M+H]$^{+}$ | 1719.7 | 6.31 |
| **R2-15** | K | R | Q | R | T | K | G | R | R | L | O | R | 1745.1 | 873.9[M+2H]$^{2+}$ | 1745.8 | 6.79 |
| **R2-16** | W | R | R | R | A | G | R | Q | R | N | R | R | 1844.1 | 615.7[M+2H]$^{3+}$<br>462.1[M+3H]$^{4+}$ | 1844.1 | 7.37 |
| **$Arg_7$** | Ac-$(Arg)_7$-$NH_2$ | | | | | | | | | | | | 1152.4 | #1152.7[M+H]$^{+}$ | 1152.7 | 5.48 |

# Matrix Assisted Laser Desorption Ionization (MALDI), Bruker Autoflex.

Analytical column: Agilent, Zorbax Eclipse XDB-C18 4.6 x250 mm (5 µm), 1 ml/min 10-90% MeCN in $H_2O$+0,1% TFA over 21.4 min (6CV).

# Appendix 4: Analytical data of the RNase P peptidomimetics

**Table 56**. Summary of the HPLC and MS analytical data of RNase P peptidomimetics [Library **P1**-Alanine scan of R2-10]. Amino acids 1 and 12 are attached to a D-Pro-L-Pro template. MWcalc.= theoretical MW calculated from molecular formula, MWobs.=MW from MS.

| Mimetic | Position | | | | | | | | | | | | MW calc | m/z ESI-MS | $MW_{obs}$ ESI-MS | $r_t$ [min] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | *7* | **8** | *9* | **10** | *11* | **12** | | | | |
| **P1-01** | A | R | L | R | A | $^{D}$R | G | K | R | N | R | R | 1686.2 | 563.5[M+3H]$^{3+}$ 423.2[M+4H]$^{4+}$ | 1687.5 | 14.24 |
| **P1-02** | W | A | L | R | A | $^{D}$R | G | K | R | N | R | R | 1716.2 | 573.0[M+3H]$^{3+}$ 430.4[M+4H]$^{4+}$ | 1716.0 | 16.95 |
| **P1-03** | W | R | A | R | A | $^{D}$R | G | K | R | N | R | R | 1759.2 | 587.8[M+3H]$^{3+}$ 439.8[M+4H]$^{4+}$ | 1760.4 | 14.57 |
| **P1-04** | W | R | L | A | A | $^{D}$R | G | K | R | N | R | R | 1716.2 | 573.4[M+3H]$^{3+}$ 430.3[M+4H]$^{4+}$ | 1717.2 | 17.37 |
| **P1-05** | W | R | L | R | A | A | G | K | R | N | R | R | 1716.2 | 573.4[M+3H]$^{3+}$ 430.1[M+4H]$^{4+}$ | 1717.2 | 16.30 |
| **P1-06** | W | R | L | R | A | $^{D}$R | A | K | R | N | R | R | 1815.2 | 605.7[M+3H]$^{3+}$ 454.7[M+4H]$^{4+}$ | 1814.8 | 15.26 |
| **P1-07** | W | R | L | R | A | $^{D}$R | G | A | R | N | R | R | 1744.1 | 582.8[M+3H]$^{3+}$ 437.3[M+4H]$^{4+}$ | 1745.4 | 16.35 |
| **P1-08** | W | R | L | R | A | $^{D}$R | G | K | A | N | R | R | 1716.2 | 573.1[M+3H]$^{3+}$ 430.4[M+4H]$^{4+}$ | 1716.3 | 16.40 |
| **P1-09** | W | R | L | R | A | $^{D}$R | G | K | R | A | R | R | 1758.1 | 587.4[M+3H]$^{3+}$ 440.4[M+4H]$^{4+}$ | 1759.2 | 16.60 |
| **P1-10** | W | R | L | R | A | $^{D}$R | G | K | R | N | A | R | 1716.2 | 573.3[M+3H]$^{3+}$ 430.2[M+4H]$^{4+}$ | 1716.6 | 16.51 |
| **P1-11** | W | R | L | R | A | $^{D}$R | G | K | R | N | R | A | 1716.2 | 573.5[M+3H]$^{3+}$ 430.4[M+4H]$^{4+}$ | 1717.5 | 18.25 |

Analytical column: Agilent, Zorbax Eclipse XDB-C18, 4.6 x 250 mm (5 µm), 6CV, 1 ml/min 10-50% MeCN in $H_2O$ +0,1%TFA.

**Table 57.** Summary of the HPLC and MS analytical data of RNase P peptidomimetics [Library-**P2**]. Amino acids 1 and 12 are attached to a D-Pro-L-Pro template. MWcalc.= theoretical MW calculated from molecular formula, MWobs.=MW from MS.

| Mimetic | Position | | | | | | | | | | | | MW calc | m/z ESI-MS | $MW_{obs}$ ESI-MS | $r_t$ [min] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | *7* | **8** | *9* | **10** | *11* | **12** | | | | |
| **P2-01** | W | R | I | R | I | $^{D}$R | G | K | R | N | R | R | 1686.2 | 563.5[M+3H]$^{3+}$ 423.2[M+4H]$^{4+}$ | 1687.5 | 14.24 |
| **P2-02** | W | A | I | R | A | $^{D}$R | G | K | R | N | R | R | 1716.2 | 573.0[M+2H]$^{3+}$ 430.4[M+3H]$^{4+}$ | 1716.0 | 16.95 |
| **P2-03** | W | R | L | R | I | $^{D}$R | G | K | R | N | R | R | 1759.2 | 587.8[M+2H]$^{3+}$ 440.9[M+3H]$^{4+}$ | 1760.4 | 14.57 |
| **P2-04** | W | R | I | R | I | $^{D}$R | N | K | R | N | R | R | 1900.3 | 634.9[M+3H]$^{3+}$ 476.4[M+4H]$^{4+}$ | 1901.4 | 19.10 |
| **P2-05** | W | R | I | R | I | $^{D}$R | N | K | R | L | R | R | 1899.3 | 634.7[M+3H]$^{3+}$ 476.0[M+4H]$^{4+}$ | 1900.0 | 20.38 |
| **P2-06** | W | R | I | R | I | $^{D}$R | N | K | R | N | Q | R | 1872.2 | 625.3[M+3H]$^{3+}$ 469.4[M+4H]$^{4+}$ | 1872.9 | 19.55 |
| **P2-07** | W | R | I | R | I | $^{D}$R | N | K | R | N | Q | A | 1787.1 | 597.2[M+3H]$^{3+}$ 448.2[M+4H]$^{4+}$ | 1788.6 | 22.88 |
| **P2-08** | W | R | I | R | I | R | N | K | R | N | R | R | 1900.3 | 635.0[M+3H]$^{3+}$ 476.5[M+4H]$^{4+}$ | 1902.0 | 18.10 |
| **P2-09** | F | R | I | R | I | $^{D}$R | N | K | R | N | R | R | 1861.3 | 621.9[M+3H]$^{3+}$ 466.8[M+4H]$^{4+}$ | 1862.7 | 19.31 |
| **P2-10** | Y | R | I | R | I | $^{D}$R | N | K | R | N | R | R | 1877.3 | 627.3[M+3H]$^{3+}$ 470.7[M+4H]$^{4+}$ | 1878.9 | 16.24 |
| **P2-11** | F | R | I | K | I | $^{D}$R | N | R | R | L | Q | A | 1747.1 | 583.4[M+3H]$^{3+}$ | 1747.2 | 26.77 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **P2-12** | W | K | I | R | I | $^{D}$R | N | K | R | N | R | R | 1872.3 | 436.8[M+4H]$^{4+}$<br>625.5[M+3H]$^{3+}$<br>469.2[M+4H]$^{4+}$ | 1872.8 | 18.48 |
| **P2-13** | W | R | I | K | I | $^{D}$R | N | K | R | N | R | R | 1872.3 | 625.6[M+3H]$^{3+}$<br>469.4[M+4H]$^{4+}$ | 1873.6 | 18.50 |
| **P2-14** | W | R | I | R | I | $^{D}$K | N | K | R | N | R | R | 1872.3 | 625.6[M+3H]$^{3+}$<br>469.4[M+4H]$^{4+}$ | 1873.6 | 17.85 |
| **P2-15** | W | R | I | R | I | $^{D}$R | N | K | K | N | R | R | 1872.3 | 625.6[M+3H]$^{3+}$<br>469.5[M+4H]$^{4+}$ | 1873.8 | 17.79 |
| **P2-16** | W | R | I | R | I | $^{D}$R | N | K | R | N | K | R | 1872.3 | 625.6[M+3H]$^{3+}$<br>469.5[M+4H]$^{4+}$ | 1873.8 | 17.82 |
| **P2-17** | W | R | I | R | I | $^{D}$R | N | K | R | N | R | K | 1872.3 | 625.7[M+3H]$^{3+}$<br>469.4[M+4H]$^{4+}$ | 1873.6 | 17.45 |
| **P2-18** | I | R | I | R | I | $^{D}$R | N | K | R | N | R | R | 1827.2 | 610.5[M+3H]$^{3+}$<br>457.9[M+4H]$^{4+}$ | 1827.6 | 17.18 |
| **P2-19** | W | Q | I | R | I | $^{D}$R | N | K | R | N | R | R | 1872.2 | 625.6[M+2H]$^{3+}$<br>469.5[M+3H]$^{4+}$ | 1873.8 | 17.51 |
| **P2-20** | W | R | I | R | I | $^{D}$R | N | Q | R | N | R | R | 1900.2 | 634.7[M+2H]$^{3+}$<br>476.5[M+3H]$^{4+}$ | 1901.1 | 17.69 |
| **P2-21** | W | C | I | R | I | $^{D}$R | N | K | R | N | C | R | 1792.2 | 896.6[M+2H]$^{2+}$<br>598.1[M+3H]$^{3+}$ | 1791.1 | 17.33 |
| **P2-22** | W | R | I | R | I | $^{D}$R | N | Q | R | N | Q | A | 1787.1 | 597.2[M+3H]$^{3+}$<br>447.7[M+4H]$^{4+}$ | 1788.6 | 22.20 |
| **P2-23** | W | R | I | R | I | $^{D}$R | N | Q | R | N | Q | R | 1872.2 | 625.6[M+3H]$^{3+}$<br>469.5[M+4H]$^{4+}$ | 1873.8 | 18.28 |
| **P2-24** | W | R | I | R | I | $^{D}$R | N | Q | Q | N | Q | A | 1759.1 | 881.1[M+2H]$^{2+}$<br>587.8[M+3H]$^{3+}$ | 1760.2 | 22.50 |

Analytical column: Agilent, Zorbax Eclipse XDB-C18, 4.6 x 250 mm (5 µm), 6CV, 1 ml/min, 10-50% MeCN in $H_2O$ +0,1%TFA.

**Table 58.** Summary of the HPLC and MS analytical data of derivatives of **P2-04**. In P2-04b amino acids 1 and 12 are attached to a D-Pro-L-Pro template. MWcalc.= theoretical MW calculated from molecular formula, MWobs.=MW from MS.

| **Mimetic** | **Position** | | | | | | | | | | | | **MW calc** | **m/z ESI-MS** | **MW$_{obs}$ ESI-MS** | **r$_t$ [min]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | *7* | **8** | *9* | **10** | *11* | **12** | | | | |
| **P2-04a** | Ac-W | R | I | R | I | $^{D}$R | N | K | R | N | R | R-$NH_2$ | 1765.1 | 589.9[M+3H]$^{3+}$<br>442.4[M+4H]$^{4+}$ | 1766.7 | 17.12 |
| **P2-04b** | $^{D}$W | $^{D}$R | $^{D}$I | $^{D}$R | $^{D}$I | R | $^{D}$N | $^{D}$K | $^{D}$R | $^{D}$N | $^{D}$R | $^{D}$R | 1900.3 | 634.2[M+3H]$^{3+}$<br>475.9[M+4H]$^{4+}$ | 1899.6 | 18.64 |

Analytical column: Agilent, Zorbax Eclipse XDB-C18, 4.6 x 250 mm (5 µm), 6CV, 1 ml/min, 10-50% MeCN in $H_2O$ +0,1%TFA.

**Table 59.** Summary of the HPLC and MS analytical data of RNase P peptidomimetics [Library-**P3**]. Amino acids 1 and 12 are attached to a D-Pro-L-Pro template. MWcalc.= theoretical MW calculated from molecular formula, MWobs.=MW from MS.

| **Mimetic** | **Position** | | | | | | | | | | | | **MW calc** | **m/z ESI-MS** | **MW$_{obs}$ ESI-MS** | **r$_t$ [min]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | *7* | **8** | *9* | **10** | *11* | **12** | | | | |
| **P3-01** | W | R | I | R | I | $^{D}$R | N | Q | K | L | R | R | 1871.3 | 936.7[M+2H]$^{2+}$<br>624.9[M+3H]$^{3+}$ | 1871.4 | 19.19 |
| **P3-02** | W | R | I | R | I | $^{D}$R | N | Q | K | I | R | R | 1871.3 | 936.5[M+2H]$^{2+}$<br>625.2[M+3H]$^{3+}$ | 1871.0 | 19.03 |
| **P3-03** | W | R | L | R | I | $^{D}$R | N | K | O | L | R | R | 1857.3 | 929.2[M+2H]$^{2+}$<br>619.7[M+3H]$^{3+}$ | 1856.4 | 18.87 |
| **P3-04** | F | R | T | R | I | $^{D}$R | N | R | R | L | E | S | 1780.1 | 890.7[M+2H]$^{2+}$<br>594.0[M+3H]$^{3+}$ | 1779.4 | 20.52 |
| **P3-05** | F | R | I | R | I | $^{D}$R | N | K | R | L | R | R | 1860.3 | 621.7[M+3H]$^{3+}$<br>466.8[M+4H]$^{4+}$ | 1862.1 | 19.24 |
| **P3-06** | W | R | I | R | I | $^{D}$R | N | K | R | L | R | I | 1856.3 | 620.4[M+3H]$^{3+}$<br>465.3[M+4H]$^{4+}$ | 1857.2 | 26.10 |
| **P3-07** | F | R | I | R | I | $^{D}$R | N | K | R | L | R | S | 1791.2 | 598.3[M+3H]$^{3+}$ | 1791.9 | 21.83 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **P3-08** | W | R | I | R | I | R | N | K | R | L | Q | A | 1786.2 | 447.8[M+4H]$^{4+}$<br>597.0[M+3H]$^{3+}$ | 1788.0 | 24.84 |
| **P3-09** | Cha | R | I | R | I | $^{D}$R | N | K | R | L | R | R | 1866.3 | 447.6[M+4H]$^{4+}$<br>623.7[M+3H]$^{3+}$ | 1868.1 | 20.33 |
| **P3-10** | W | R | I | R | I | $^{D}$R | N | Q | R | L | Q | A | 1786.2 | 468.2[M+4H]$^{4+}$<br>596.9[M+3H]$^{3+}$<br>447.7[M+4H]$^{4+}$ | 1787.7 | 28.08 |

Analytical column: Agilent, Zorbax Eclipse XDB-C18, 4.6 x250 mm (5 μm), 6CV, 1 ml/min, 10-50% MeCN in $H_2O$ +0,1%TFA.

# Appendix 5: Antimicrobial peptides-derivatives of M18A11

The growing problem of resistance to established antibiotics has stimulated intense interest in the development of novel antimicrobial agents with new modes of action. One emerging class of antibiotics is based on naturally occurring cationic peptides, which have been discovered in many animal species, where they are thought to constitute ancient evolutionary elements of immunity.[213,214] These include disulfide-bridged β-hairpin and β-sheet peptides (for example, the protegrins, tachyplesins,[215] and defensins[216]), amphipatic α-helical peptides (for example, cecropins, dermaseptins, magainins, and mellitins),[217] as well as other linear and loop-structured peptides.[218] The primary site of interaction of these cationic peptides is the microbial cell membrane.[214,218] Upon exposure to these agents, the cell membrane undergoes permeabilization, which leads to rapid cell death.

Structural studies by NMR spectroscopy have shown that protegrin[219] and tachyplesin[220] adopt well-defined β-hairpin structures in water, as a result of the constraining effect of the two disulfide cross-links. In protegrin analogues that lack one or both of the disulfide bridges, the stability of the β-harpin conformation is diminished and the membranolytic activity is reduced.[193] Similar observations were made for tachyplesin,[221] which indicates that the β-hairpin structure plays an important role in the activity of these molecules.

The starting point for the antimicrobial peptidomimetic design in our group was the naturally occurring antimicrobial peptide protegrin I (PG-I), one of the family of five homologous peptides first isolated from porcine leucocytes.[222] PG-I is an 18-amino acid peptide containing two disulfide cross-links (Cys6-15 and Cys8-13). These disulfide bridges constrain the peptide into a β-hairpin conformation, with a β-turn formed by residues 9-12. PG-I typically kills susceptible microorganisms within minutes,[208] consistent with a bactericidal effect caused by disruption of the cell membrane.[221,223]

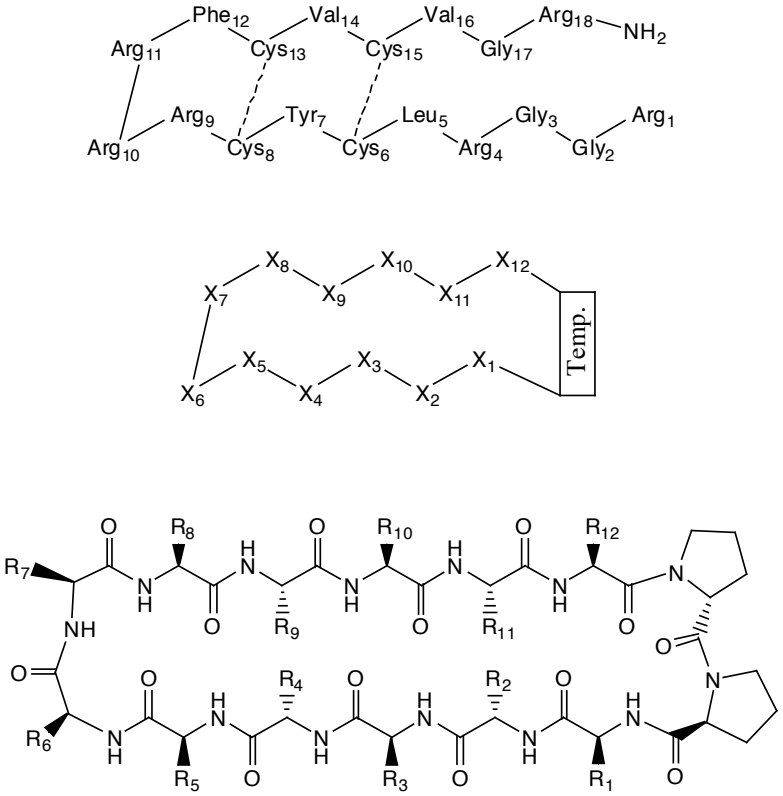



**Figure 79.** Naturally occurring cationic β-hairpin antimicrobial peptide, PG-I (top), served as a starting point for the design of template-bound antimicrobial peptidomimetics, shown below.

The design of PG-I mimetics was performed by grafting the hairpin sequence onto an D-Pro-L-Pro template whose function is to restrain the peptide loop termini in a β-hairpin geometry (Figure 79). The extensive SAR studies conducted with the 14-mer peptidomimetics led to the development of the very potent antimicrobial agents (**D** and **M18A11,** Table 60) with high selectivity and low haemolytic activity in comparison to PG-I and other first generation peptides (e.g. **A**).[224,225]

**Table 60.** Antimicrobial and haemolytic activities of antimicrobial peptidomimetics derived from protegrin-1. The values for protegrin-1 and mimetics A and B are taken from Shankaramma et al.,[225] and the others are taken from PhD Thesis of N. Srinivas.[224] nd=not determined, % haemolysis at a peptide concentration of 100 μg/ml.

| **Peptide** | **Sequence** | **MIC[μg/ml]** | | | | | | **% H** |
|---|---|---|---|---|---|---|---|---|
| | | *E. coli* | *P. aeruginosa* | | *S. aureus* | | *C. albicans* | hRBC |
| | | 25922 | 27853 | PAO1 | 25923 | 29213 | | |
| **PG-I** | Protegrin-1 | 3 | 3 | nd | 6 | nd | 6 | 37 |
| **A** | LRLQYRRFQYRVpP | 6 | 12 | nd | 12 | nd | 6 | 27 |
| **B** | LRLKKRRWKYRVpP | 12 | 6 | nd | 12 | nd | 12 | 1 |
| **C** | RWLKKQRWKYYRpP | 16 | 0.25 | nd | >64 | >64 | nd | 0 |
| **D** | TWLKKRRWKKVKpP | 64 | 0.03 | 0.06 | 32 | >64 | nd | 0 |
| **M18A11** | TWLKKRRWKKAKpP | 64 | 0.015 | 0.004 | >64 | >64 | nd | 0 |

The peptidomimetics from library KP1 (Table 61) and KP2 (Table 63) were designed based on the sequence of the most potent antimicrobial peptide produced so far, **M18A11**, with the goal to investigate the effect of changing the nature of the positively charged residues. In these libraries the arginine residues present in the sequence were mutated to lysines followed by the single/mutiple mutations of Arg/Lys to the unnatural amino acids ornithine (O) and diaminobutyric acid (Dab). It was hypothesised that introduction of these unnatural amino acids will increase the stability of the peptides in biological settings, due to their resistance to proteolytic degradation.

**Table 61.** Summary of the HPLC and MS analytical data of antimicrobial peptidomimetics in library KP1. Amino acids 1 and 12 are attached to a D-Pro-L-Pro template. MWcalc.= theoretical MW calculated from molecular formula, MWobs.=MW from MS. O=Ornithine

| **Mimetic** | **Position** | | | | | | | | | | | | **MW calc** | **m/z ESI-MS** | **$MW_{obs}$ ESI-MS** | **$r_t$ [min]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | *7* | **8** | *9* | **10** | *11* | **12** | | | | |
| **KP1-01** | T | W | L | K | K | R | R | W | K | K | A | K | 1805.3 | 602.7[M+3H]$^{3+}$<br>452.6[M+4H]$^{4+}$ | 1805.1 | 18.63 |
| **KP1-02** | T | W | L | K | K | K | R | W | K | K | A | K | 1777.5 | 593.7[M+3H]$^{3+}$<br>445.3[M+4H]$^{4+}$ | 1778.1 | 18.25 |
| **KP1-03** | T | W | L | K | K | R | K | W | K | K | A | K | 1777.5 | 593.7[M+3H]$^{3+}$<br>445.5[M+4H]$^{4+}$ | 1778.1 | 18.49 |
| **KP1-04** | T | W | L | K | K | K | K | W | K | K | A | K | 1749.2 | 583.9[M+3H]$^{3+}$<br>438.6[M+4H]$^{4+}$ | 1748.7 | 18.05 |
| **KP1-05** | T | W | L | O | K | K | K | W | K | K | A | K | 1735.1 | 579.8[M+3H]$^{3+}$<br>435.1[M+4H]$^{4+}$ | 1736.4 | 18.06 |
| **KP1-06** | T | W | L | K | O | K | K | W | K | K | A | K | 1735.1 | 579.7[M+3H]$^{3+}$<br>435.1[M+4H]$^{4+}$ | 1736.1 | 18.35 |
| **KP1-07** | T | W | L | K | K | K | K | W | O | K | A | K | 1735.1 | 579.8[M+3H]$^{3+}$<br>435.1[M+4H]$^{4+}$ | 1736.4 | 18.21 |
| **KP1-08** | T | W | L | K | K | K | K | W | K | O | A | K | 1735.1 | 579.8[M+3H]$^{3+}$<br>434.9[M+4H]$^{4+}$ | 1736.4 | 18.00 |
| **KP1-09** | T | W | L | K | K | K | K | W | K | K | A | O | 1735.1 | 579.6[M+3H]$^{3+}$<br>435.0[M+4H]$^{4+}$ | 1735.8 | 18.02 |
| **KP1-10** | T | W | L | K | K | O | O | W | K | K | A | K | 1720.9 | 575.0[M+3H]$^{3+}$<br>431.4[M+4H]$^{4+}$ | 1722.0 | 18.09 |
| **KP1-11** | T | W | L | O | O | K | K | W | O | O | A | O | 1678.3 | 561.0[M+3H]$^{3+}$<br>420.8[M+4H]$^{4+}$ | 1680.0 | 18.08 |
| **KP1-12** | T | W | L | O | O | O | O | W | O | O | A | O | 1650.0 | 551.6[M+3H]$^{3+}$<br>413.8[M+4H]$^{4+}$ | 1651.8 | 18.01 |

Analytical column: Agilent, Zorbax Eclipse XDB-C18, 4.6 x 250 mm (5 μm), 6CV, 1 ml/min, 10-50% MeCN in $H_2O$ +0,1%TFA

**Table 62.** Antimicrobial activity (MIC) measured for peptidomimetics from the KP1 library.

| **Mimetic** | ***P. putida* 2442**<br>**MIC [μg/ml]** | ***P. aeruginosa PAOI***<br>**MIC [μg/ml]** |
|---|---|---|
| **KP1-01** | 0.25 | 0.006 |
| **KP1-02** | 0.12 | 0.002 |
| **KP1-03** | 0.12 | 0.003 |
| **KP1-04** | 0.12 | 0.002 |
| **KP1-05** | 0.25 | 0.003 |

| | | |
|---|---|---|
| **KP1-06** | 2 | 0.002 |
| **KP1-07** | 0.03 | 0.006 |
| **KP1-08** | 0.03 | 0.002 |
| **KP1-09** | 0.25 | 0.006 |
| **KP1-10** | 0.004 | 0.003 |
| **KP1-11** | 1 | 0.002 |
| **KP1-12** | 0.75 | 0.002 |

**Table 63.** Summary of the HPLC and MS analytical data of antimicrobial peptidomimetics in library KP2. Amino acids 1 and 12 are attached to a D-Pro-L-Pro template. MWcalc.= theoretical MW calculated from molecular formula, MWobs.=MW from MS. Dab=diaminobutyric acid, O=ornithine

| **Mimetic** | **Position** | | | | | | | | | | | | **MW calc** | **m/z ESI-MS** | **$MW_{obs}$ ESI-MS** | **$r_t$ [min]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | *7* | **8** | *9* | **10** | *11* | **12** | | | | |
| **KP2-01** | T | W | L | O | O | O | O | W | O | O | A | O | 1650.0 | 551.6[M+3H]$^{3+}$ 413.8[M+4H]$^{4+}$ | 1651.8 | 18.01 |
| **KP2-02** | T | W | L | Dab | O | O | O | W | O | O | A | O | 1636.2 | 547.0[M+3H]$^{3+}$ 410.7[M+4H]$^{4+}$ | 1638.0 | 16.08 |
| **KP2-03** | T | W | L | O | Dab | O | O | W | O | O | A | O | 1636.2 | 546.9[M+3H]$^{3+}$ 410.6[M+4H]$^{4+}$ | 1637.7 | 15.87 |
| **KP2-04** | T | W | L | O | O | Dab | O | W | O | O | A | O | 1636.2 | 547.0[M+3H]$^{3+}$ 410.7[M+4H]$^{4+}$ | 1638.0 | 15.83 |
| **KP2-05** | T | W | L | O | O | O | Dab | W | O | O | A | O | 1636.2 | 547.1[M+3H]$^{3+}$ 410.6[M+4H]$^{4+}$ | 1638.3 | 16.03 |
| **KP2-06** | T | W | L | O | O | O | O | W | Dab | O | A | O | 1636.2 | 547.0[M+3H]$^{3+}$ 410.6[M+4H]$^{4+}$ | 1638.0 | 16.09 |
| **KP2-07** | T | W | L | O | O | O | O | W | O | Dab | A | O | 1636.2 | 546.7[M+3H]$^{3+}$ 410.8[M+4H]$^{4+}$ | 1637.1 | 16.01 |
| **KP2-08** | T | W | L | O | O | O | O | W | O | O | A | Dab | 1636.2 | 547.0[M+3H]$^{3+}$ 410.7[M+4H]$^{4+}$ | 1638.0 | 16.10 |
| **KP2-09** | T | W | L | Dab | Dab | Dab | Dab | W | Dab | Dab | A | Dab | 1552.2 | 518.9[M+3H]$^{3+}$ 388.1[M+4H]$^{4+}$ | 1553.7 | 16.10 |

Analytical column: Agilent, Zorbax Eclipse XDB-C18, 4.6 x 250 mm (5 μm), 6CV, 1 ml/min, 10-50% MeCN in $H_2O$ +0,1%TFA

**Table 64.** Antimicrobial activity (MIC) measured for peptidomimetics from the KP2 library.

| **Mimetic** | ***P. putida* 2442 MIC [μg/ml]** | ***P. aeruginosa PAOI* MIC [μg/ml]** |
|---|---|---|
| **KP2-01** | 0.25 | 0.006 |
| **KP2-02** | 0.12 | 0.001 |
| **KP2-03** | 16 | 0.02 |
| **KP2-04** | 0.33 | 0.005 |
| **KP2-05** | 0.12 | 0.001 |
| **KP2-06** | 1 | 0.001 |
| **KP2-07** | 0.02 | 0.001 |
| **KP2-08** | 0.09 | 0.001 |
| **KP2-09** | 4 | 0.1 |

**Table 65.** Trypsin digestion assay. Shown are the half-lives of the peptides in the trypsin digestion assay (for detail see paragraph 2.2.8.1).

| **Peptide** | **Half-life $\tau_{1/2}$ [min]** |
|---|---|
| **KP1-01** | <5 min |
| **KP1-12** | >13 days (after 3 days, 90%left; after 13 days, 65%left) |
| **KP2-09** | 5 days |

The results of the antimicrobial activities presented in Table 62 and Table 64 indicate that replacement of the arginine and lysine residues does not bring much change in the antimicrobial activity against *P. aeruginosa*, while larger changes were observed for *P. putida*. Only the peptide containing only Dab residues, **KP2-09**, shows a 20-fold reduced inhibitory potency against *P. aeruginosa* in comparison with **KP1-01**. Some of the peptides exhibit an increase of 2- to 4-fold in inhibitory activity, which might be due to their improved stability towards proteolysis. The lead peptide, **M18A11** (called here **KP1-01**) is digested completely in 30 min whereas the two peptides having exclusively ornithine residues (**KP1-12**), or diaminobutyric acid (**KP2-09**) are stable in the assay conditions for at least a few days (Table 65).

The first part of the KP3 library (**KP3-02**—**KP3-12**) represents an alanine scan prepared for **M18A11** (**KP3-01**), and performed in order to investigate the importance of the single amino acid residues for the antimicrobial activity. The second part of the library, with peptides **KP3-13**—**KP3-34**, represents single mutations of the two tryptophan residues present in the sequence of **M18A11** peptide. It was performed to assess the tolerance to the replacement of these critical amino acids for other hydrophobic or aromatic side chains.

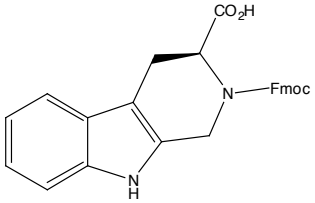
**Table 66.** Summary of the HPLC and MS analytical data of antimicrobial peptidomimetics [Alanine scan and mutations of tryptophan in **M18A11**]. Amino acids 1 and 12 are attached to a D-Pro-L-Pro template. MWcalc.= theoretical MW calculated from molecular formula, MWobs.=MW from MS.

| **Mimetic** | **Position** | | | | | | | | | | | | **MW calc** | **m/z ESI-MS** | **$MW_{obs}$ ESI-MS** | **$r_t$ [min]** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | *2* | **3** | *4* | **5** | *6* | *7* | **8** | *9* | **10** | *11* | **12** | | | | |
| **KP3-01** | T | W | L | K | K | R | R | W | K | K | A | K | 1805.3 | 602.7[M+3H]$^{3+}$<br>452.6[M+4H]$^{4+}$ | 1805.1 | 17.07 |
| **KP3-02** | A | W | L | K | K | R | R | W | K | K | A | K | 1775.5 | 592.9[M+3H]$^{3+}$<br>445.1[M+4H]$^{4+}$ | 1775.7 | 17.66 |
| **KP3-03** | T | A | L | K | K | R | R | W | K | K | A | K | 1690.1 | 564.5[M+3H]$^{3+}$<br>423.8[M+4H]$^{4+}$ | 1690.5 | 15.35 |
| **KP3-04** | T | W | A | K | K | R | R | W | K | K | A | K | 1763.2 | 589.0[M+3H]$^{3+}$<br>442.1[M+4H]$^{4+}$ | 1764.0 | 14.38 |
| **KP3-05** | T | W | L | A | K | R | R | W | K | K | A | K | 1748.2 | 583.8[M+3H]$^{3+}$<br>438.1[M+4H]$^{4+}$ | 1748.4 | 17.59 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **KP3-06** | T | W | L | K | A | R | R | W | K | K | A | K | 1748.2 | 583.8[M+3H]$^{3+}$ 438.1[M+4H]$^{4+}$ | 1748.4 | 17.90 |
| **KP3-07** | T | W | L | K | K | A | R | W | K | K | A | K | 1720.2 | 574.7[M+3H]$^{3+}$ 431.3[M+4H]$^{4+}$ | 1721.1 | 16.54 |
| **KP3-08** | T | W | L | K | K | R | A | W | K | K | A | K | 1720.2 | 574.8[M+3H]$^{3+}$ 431.3[M+4H]$^{4+}$ | 1721.4 | 17.66 |
| **KP3-09** | T | W | L | K | K | R | R | A | K | K | A | K | 1690.1 | 564.8[M+3H]$^{3+}$ 423.8[M+4H]$^{4+}$ | 1691.4 | 15.57 |
| **KP3-10** | T | W | L | K | K | R | R | W | A | K | A | K | 1748.2 | 583.9[M+3H]$^{3+}$ 438.1[M+4H]$^{4+}$ | 1748.7 | 17.94 |
| **KP3-11** | T | W | L | K | K | R | R | W | K | A | A | K | 1748.2 | 583.4[M+3H]$^{3+}$ 438.1[M+4H]$^{4+}$ | 1747.2 | 16.86 |
| **KP3-12** | T | W | L | K | K | R | R | W | K | K | A | A | 1748.2 | 583.9[M+3H]$^{3+}$ 438.1[M+4H]$^{4+}$ | 1748.7 | 18.18 |
| **KP3-13** | T | F | L | K | K | R | R | W | K | K | A | K | 1766.2 | 590.0[M+3H]$^{3+}$ 442.6[M+4H]$^{4+}$ | 1767.0 | 17.95 |
| **KP3-14** | T | Y | L | K | K | R | R | W | K | K | A | K | 1782.2 | 595.4[M+3H]$^{3+}$ 446.8[M+4H]$^{4+}$ | 1783.2 | 16.35 |
| **KP3-15** | T | L | L | K | K | R | R | W | K | K | A | K | 1732.2 | 578.7[M+3H]$^{3+}$ 434.2[M+4H]$^{4+}$ | 1733.1 | 17.07 |
| **KP3-16** | T | Cha | L | K | K | R | R | W | K | K | A | K | 1772.3 | 592.1[M+3H]$^{3+}$ 444.4[M+4H]$^{4+}$ | 1773.3 | 19.29 |
| **KP3-17** | T | 1-Nal | L | K | K | R | R | W | K | K | A | K | 1816.3 | 606.4[M+3H]$^{3+}$ 455.3[M+4H]$^{4+}$ | 1816.2 | 19.23 |
| **KP3-18** | T | 2-Nal | L | K | K | R | R | W | K | K | A | K | 1816.3 | 606.7[M+3H]$^{3+}$ 455.4[M+4H]$^{4+}$ | 1817.1 | 19.76 |
| **KP3-19** | T | Hfe | L | K | K | R | R | W | K | K | A | K | 1780.3 | 594.7[M+3H]$^{3+}$ 446.1[M+4H]$^{4+}$ | 1781.1 | 18.78 |
| **KP3-20** | T | 3Cl-F | L | K | K | R | R | W | K | K | A | K | 1800.7 | 601.4[M+3H]$^{3+}$ 451.5[M+4H]$^{4+}$ | 1801.2 | 18.74 |
| **KP3-21** | T | 4Cl-F | L | K | K | R | R | W | K | K | A | K | 1800.7 | 601.7[M+3H]$^{3+}$ 451.2[M+4H]$^{4+}$ | 1802.1 | 18.31 |
| **KP3-22** | T | Bip | L | K | K | R | R | W | K | K | A | K | 1842.4 | 615.4[M+3H]$^{3+}$ 461.9[M+4H]$^{4+}$ | 1843.2 | 20.51 |
| **KP3-23** | T | Tpi | L | K | K | R | R | W | K | K | A | K | 1817.1 | - | - | - |
| **KP3-24** | T | W | L | K | K | R | R | F | K | K | A | K | 1766.2 | 590.1[M+3H]$^{3+}$ 442.6[M+4H]$^{4+}$ | 1767.3 | 17.16 |
| **KP3-25** | T | W | L | K | K | R | R | Y | K | K | A | K | 1782.2 | 595.4[M+3H]$^{3+}$ 446.8[M+4H]$^{4+}$ | 1783.2 | 16.25 |
| **KP3-26** | T | W | L | K | K | R | R | L | K | K | A | K | 1732.2 | 578.8[M+3H]$^{3+}$ 434.2[M+4H]$^{4+}$ | 1733.4 | 16.99 |
| **KP3-27** | T | W | L | K | K | R | R | Cha | K | K | A | K | 1772.3 | 592.1[M+3H]$^{3+}$ 444.2[M+4H]$^{4+}$ | 1773.3 | 17.78 |
| **KP3-28** | T | W | L | K | K | R | R | 1-Nal | K | K | A | K | 1816.3 | 606.8[M+3H]$^{3+}$ 455.3[M+4H]$^{4+}$ | 1817.4 | 17.95 |
| **KP3-29** | T | W | L | K | K | R | R | 2-Nal | K | K | A | K | 1816.3 | 606.7[M+3H]$^{3+}$ 455.2[M+4H]$^{4+}$ | 1817.1 | 18.24 |
| **KP3-30** | T | W | L | K | K | R | R | Hfe | K | K | A | K | 1780.3 | 594.7[M+3H]$^{3+}$ 446.2[M+4H]$^{4+}$ | 1781.1 | 17.45 |
| **KP3-31** | T | W | L | K | K | R | R | 3Cl-F | K | K | A | K | 1800.7 | 601.7[M+3H]$^{3+}$ 451.2[M+4H]$^{4+}$ | 1802.1 | 17.88 |
| **KP3-32** | T | W | L | K | K | R | R | 4Cl-F | K | K | A | K | 1800.7 | 601.0[M+3H]$^{3+}$ 451.5[M+4H]$^{4+}$ | 1800.0 | 17.69 |
| **KP3-33** | T | W | L | K | K | R | R | Bip | K | K | A | K | 1842.4 | 615.3[M+3H]$^{3+}$ 461.9[M+4H]$^{4+}$ | 1842.9 | 19.20 |
| **KP3-34** | T | W | L | K | K | R | R | Tpi | K | K | A | K | 1817.1 | - | - | - |

Cha=Cyclohexylalanine
3Cl-F=3-chloro-phenylalanine, 4Cl-F=4-chloro-phenylalanine
Bip = 4-Phenyl-phenylalanine
Hfe = Homophenylalanine
1(2)-Nal=1(2)-Naphthylalanine

Analytical column: Agilent, Zorbax Eclipse XDB-C18, 4.6 x 250 mm (5 µm), 6CV, 1 ml/min, 10-50% MeCN in $H_2O$ +0,1%TFA




Fmoc-L-1,2,3,4-tetrahydronorharman-3-carboxylic acid
Fmoc-L-Tpi-OH

The syntheses of peptides containing 1,2,3,4-tetrahydronorharman-3-carboxylic acid (**KP3-23** and **KP3-34**) failed.

**Table 67.** Antimicrobial activity (MIC) for the peptides from the KP3 library.

| **Mimetic** | ***P. aeruginosa PAOI*** **MIC [μg/ml]** |
|---|---|
| **KP3-01** | 0.02 |
| **KP3-02** | 0.02 |
| **KP3-03** | >16 |
| **KP3-04** | 0.09 |
| **KP3-05** | 0.05 |
| **KP3-06** | 0.09 |
| **KP3-07** | 0.02 |
| **KP3-08** | 0.04 |
| **KP3-09** | >16 |
| **KP3-10** | 0.04 |
| **KP3-11** | 0.04 |
| **KP3-12** | 0.02 |
| **KP3-13** | 0.04 |
| **KP3-14** | 0.04 |
| **KP3-15** | 1 |
| **KP3-16** | 0.25 |
| **KP3-17** | 0.2 |
| **KP3-18** | 0.09 |
| **KP3-19** | 3 |
| **KP3-20** | 0.03 |
| **KP3-21** | 0.04 |
| **KP3-22** | 0.04 |
| **KP3-24** | 0.09 |
| **KP3-25** | 0.25 |
| **KP3-26** | 0.2 |
| **KP3-27** | 0.08 |
| **KP3-28** | 0.05 |
| **KP3-29** | 0.005 |
| **KP3-30** | 0.04 |
| **KP3-31** | 0.2 |
| **KP3-32** | 0.05 |
| **KP3-33** | 0.8 |

The results of the antimicrobial assays of the alanine-scanning library prove the importance of the two tryptophan residues in the sequence. Therefore, the single replacements of these critical amino acids were performed using natural (Phe, Leu, Tyr) and unnatural amino acids (homophenylalanine; cyclohexylalanine; 3-, 4-chloro-phenyalanine; 4-phenyl-phenylalanine; 1-,2-naphtylalanine). Overall the changes are were tolerated, but generally with a slight decrease in potency. A significant decrease is

noticed only for the Trp3Leu and Trp3Hfe mutants (40 and 80-fold). More than 10-fold reduction in activity is observed for the mutants of Trp3Cha, Trp8Tyr and Trp8Bip. Only one mutation, of Trp8 to 2-Naphthylalanine, brings an increase in the antimicrobial activity (4-fold, MIC=0.005 vs 0.02 μg/ml) in comparison to **M18A11** (**KP3-01**).

# Appendix 6: NMR data for BIV/HIV Tat and Rev peptidomimetics

The peptide backbone NH, C(α)H and side chain resonances for all the studied cyclic peptides were assigned unambiguously.

**Table 68.** $^{1}$H-NMR assignment of **L1-07**. Chemical shifts obtained relative to the internal TSP standard at 298K and pH=2.3 in $H_2O/D_2O$ (9:1).

| L1-07 | NH | α | β | Others |
|---|---|---|---|---|
| Arg$^{1}$ | 7.754 | 4.676 | 1.853 | $CH_2$(γ) 1.622; $CH_2$(δ) 3.214; NH(ε) 7.282 |
| Thr$^{2}$ | 8.501 | 5.009 | 3.923 | $CH_3$(γ) 1.041 |
| Arg$^{3}$ | 9.086 | 4.722 | 1.675 1.830 | $CH_2$(γ) 1.532; $CH_2$(δ) 3.159; NH(ε) 7.242 |
| Thr$^{4}$ | 8.641 | 4.933 | 4.029 | $CH_3$(γ) 1.073 |
| Arg$^{5}$ | 8.726 | 4.507 | 1.588 1.747 | $CH_2$(γ) 1.476, 1.589; $CH_2$(δ) 3.160 |
| Gly$^{6}$ | 9.139 | 3.697, 3.996 | - | - |
| Lys$^{7}$ | 8.913 | 4.231 | 1.751 2.016 | $CH_2$(γ) 1.427, 1.464; $CH_2$(δ) 1.673; $CH_2$(ε) 2.985 |
| Arg$^{8}$ | 7.905 | 4.502 | 1.863 | $CH_2$(γ) 1.591, 1.687; $CH_2$(δ) 3.209 |
| Arg$^{9}$ | 8.531 | 4.515 | 1.579 1.707 | $CH_2$(γ) 1.472, 1.582; $CH_2$(δ) 3.109 |
| Ile$^{10}$ | 8.958 | 4.443 | 1.758 | $CH_2$(γ) 1.031, 1.320; $CH_3$($\gamma^2$) 0.831 $CH_3$($\delta^1$) 0.753 |
| Arg$^{11}$ | 8.681 | 4.691 | 1.736 1.829 | $CH_2$(γ) 1.548, 1.637; $CH_2$(δ) 3.158; NH(ε) 8.716 |
| Val$^{12}$ | 8.716 | 4.455 | 1.938 | $CH_3$(γ) 0.816, 0.879 |
| Pro$^{13}$ | - | 4.734 | 1.914 2.293 | $CH_2$(γ) 2.004, 2.124; $CH_2$(δ) 3.479, 3.849 |
| Pro$^{14}$ | - | 4.562 | 2.124 2.245 | $CH_2$(γ)1.890, 2.107; $CH_2$(δ) 3.744, 3.975 |

**Table 69.** $^{1}$H-NMR assignment of **L1-43**. Chemical shifts obtained relative to the internal TSP standard at 298K and pH=2.3 in $H_2O/D_2O$ (9:1).

| L1-43 | NH | α | β | Others |
|---|---|---|---|---|
| Arg$^{1}$ | 7.734 | 4.657 | 1.833 | $CH_2$(γ) 1.611; $CH_2$(δ) 3.207; NH(ε) |
| Val$^{2}$ | 8.453 | 4.725 | 1.860 | $CH_3$(γ) 0.771, 0.852 |
| Arg$^{3}$ | 9.120 | 4.750 | 1.667, 1.841 | $CH_2$(γ) 1.519, 1.575; $CH_2$(δ) 3.146, 3.178; |
| Thr$^{4}$ | 8.633 | 4.953 | 4.027 | $CH_3$(γ) 1.073 |
| Arg$^{5}$ | 8.704 | 4.517 | 1.586, 1.744 | $CH_2$(γ) 1.473; $CH_2$(δ) 3.171 |
| Gly$^{6}$ | 9.150 | 3.690, 3.993 | - | - |
| Lys$^{7}$ | 8.926 | 4.231 | 1.757, 2.031 | $CH_2$(γ) 1.425, 1.452; $CH_2$(δ) 1.671; $CH_2$(ε) 2.990 |
| Arg$^{8}$ | 4.490 | 4.491 | 1.862 | $CH_2$(γ) 1.593, 1.697; $CH_2$(δ) 3.211 |
| Arg$^{9}$ | 8.507 | 4.801 | 1.568, 1.680 | $CH_2$(γ) 1.464; $CH_2$(δ) 3.111 |
| Ile$^{10}$ | 8.988 | 4.416 | 1.746 | $CH_2$(γ) 1.033, 1.338; $CH_3$($\gamma^2$) 0.834 $CH_3$($\delta^1$) 0.759 |
| Orn$^{11}$ | 8.734 | 4.690 | 1.716, 1.995 | $CH_2$(γ) 1.715; $CH_2$(δ) 2.914, 3.002; NH(ε) |
| Val$^{12}$ | 8.832 | 4.490 | 1.937 | $CH_3$(γ) 0.833, 0.885 |
| Pro$^{13}$ | - | 4.755 | 1.926, 2.300 | $CH_2$(γ) 2.017, 2.137; $CH_2$(δ) 3.517, 3.870 |
| Pro$^{14}$ | - | 4.568 | 2.141, 2.234 | $CH_2$(γ)1.888, 2.110; $CH_2$(δ) 3.739, 3.957 |

**Table 70.** $^1$H-NMR assignment of **L2-15**. Chemical shifts obtained relative to the internal TSP standard at 298K and pH=2.3 in $H_2O/D_2O$ (9:1).

| **L2-15** | **NH** | **α** | **β** | **Others** |
|---|---|---|---|---|
| Arg$^1$ | 7.751 | 4.604 | 1.809 | $CH_2$(γ) 1.603; $CH_2$(δ) 3.204; NH(ε) 7.272 |
| Val$^2$ | 8.461 | 4.733 | 1.867 | $CH_3$(γ) 0.844, $CH_3$(γ) 0.784 |
| Arg$^3$ | 8.971 | 4.737 | 1.624 1.846 | $CH_2$(γ) 1.555, 1.524; |
| Thr$^4$ | 8.585 | 5.092 | 3.928 | $CH_3$(γ) 1.002 |
| Arg$^5$ | 8.976 | 4.549 | 1.733, 1.623 | $CH_2$(γ) 1.622, 1.439; |
| Lys$^6$ | 9.590 | 3.900 | 2,012, 1.804 | $CH_2$(γ) 1.412, $CH_2$(ε) 2.993, |
| Gly$^7$ | 8.477 | 4.125, 3.564 | - | - |
| Arg$^8$ | 7.840 | 4.617 | 1.845, 1.785 | $CH_2$(γ) 1.639, 1.588; |
| Arg$^9$ | 8.630 | 4.690 | 1.706 1.582 | $CH_2$(γ) 1.581, 1.456; $CH_2$(δ) 3.110; NH(ε) 7.166 |
| Ile$^{10}$ | 9.072 | 4.444 | 1.788 | $CH_3$(γ) 1.256, 1.061 $CH_3$($\gamma^2$) 0.841 $CH_3$($\delta^1$) 0.782 |
| Arg$^{11}$ | 8.582 | 4.822 | 1.828 1.669 | $CH_2$(γ) 1.568, 1.484; |
| Ile$^{12}$ | 8.750 | 4.552 | 1.774 | $CH_3$(γ) 1.387, 1.060 $CH_3$($\gamma^2$) 0.854 |
| Pro$^{13}$ | - | 4.733 | 1.909 2.287 | $CH_2$(γ) 2.006, 2.121; $CH_2$(δ) 3.521, 3.860 |
| Pro$^{14}$ | - | 4.547 | 2.119, 2.233 | $CH_2$(γ)1.880, 2.104; $CH_2$(δ) 3.738, 3.970 |

**Table 71.** $^1$H-NMR assignment of **L4-01**. Chemical shifts obtained relative to the internal TSP standard at 298K and pH=2.3 in $H_2O/D_2O$ (9:1).

| **L4-01** | **NH** | **α** | **β** | **Others** |
|---|---|---|---|---|
| Arg$^1$ | 7.658 | 4.667 | 1.805 | $CH_2$(γ) 1.589; $CH_2$(δ) 3.200; NH(ε) 7.276 |
| Val$^2$ | 8.536 | 4.478 | 1.887 | $CH_3$($\gamma^2$) 0.809, 0.861 |
| Arg$^3$ | 8.865 | 4.754 | 1.727, 1.846 | $CH_2$(γ) 1.511, 1.659; $CH_2$(δ) 3.176; NH(ε) 7.256 |
| Cys$^4$ | 9.129 | 5.616 | 2.580, 3.046 | - |
| Arg$^5$ | 8.857 | 4.673 | 1.763, 1.839 | $CH_2$(γ) 1.605; $CH_2$(δ) 3.201; NH(ε) 7.197 |
| Thr$^6$ | 8.797 | 4.869 | 4.003 | $CH_3$(γ) 1.033 |
| Arg$^7$ | 8.529 | 4.516 | 1.685, 1.784 | $CH_2$(γ) 1.516, 1.611; $CH_2$(δ); NH(ε) |
| Gly$^8$ | 9.161 | 3.734, 3.968 | - | - |
| Lys$^9$ | 8.862 | 4.159 | 1.788, 2.020 | $CH_2$(γ) 1.397, 1.440; $CH_2$(δ) 1.673; $CH_2$(ε) 2.985; NH(ε) 7.276 |
| Arg$^{10}$ | 7.847 | 4.560 | 1.855 | $CH_2$(γ) 1.580, 1.662; $CH_2$(δ) 3.206; $NH_2$(ε) 7.590 |
| Arg$^{11}$ | 8.668 | 4.840 | 1.636, 1.747 | $CH_2$(γ) 1.427; $CH_2$(δ) 3.092; $NH_2$(ε) 7.159 |
| Arg$^{12}$ | 8.729 | 4.703 | 1.743, 1.827 | $CH_2$(γ) 1.561, 1.632; $CH_2$(δ) 3.186; NH(ε) |
| Cys13 | 9.027 | 5.516 | 2.634, 2.988 | - |
| Ala$^{14}$ | 8.831 | 4.615 | 1.366 | - |
| Arg$^{15}$ | 8.724 | 4.667 | 1.709 | $CH_2$(γ) 1.500, 1.562; $CH_2$(δ) 3.183 |
| Val$^{16}$ | 8.828 | 4.484 | 1.955 | $CH_3$($\gamma^2$) 0.886 |
| Pro$^{17}$ | - | 4.759 | 1.911, 2.292 | $CH_2$(γ) 2.015, 2.136; $CH_2$(δ) 3.529, 3.877 |
| Pro$^{18}$ | - | 4.559 | 2.130, 2.238 | $CH_2$(γ) 1.876, 2.102; $CH_2$(δ) 3.747, 3.975 |

**Table 72.** $^1$H-NMR assignment of **L4-04**. Chemical shifts obtained relative to the internal TSP standard at 298K and pH=2.3 in $H_2O/D_2O$ (9:1).

| **L4-04** | **NH** | **α** | **β** | **Others** |
|---|---|---|---|---|
| $Arg^1$ | 7.747 | 4.609 | 1.826 | $CH_2(\gamma)$ 1.596; $CH_2(\delta)$ 3.199; $NH(\varepsilon)$ 7.280 |
| $Val^2$ | 8.488 | 4.687 | 1.880 | $CH_3(\gamma^2)$ 0.832, |
| $Arg^3$ | 8.905 | 4.725 | 1.658, 1.805 | $CH_2(\gamma)$ 1.479, 1.656; $CH_2(\delta)$ 3.156; $NH(\varepsilon)$ 7.230 |
| $Cys^4$ | 9.171 | 5.620 | 2.610, 3.015 | - |
| $Arg^5$ | 8.798 | 4.646 | 1.789 | $CH_2(\gamma)$ 1.586; $CH_2(\delta)$ 3.134, 3.194; $NH(\varepsilon)$ 7.230 |
| $Lys^6$ | 8.827 | 4.947 | 1.636, 1.702 | $CH_2(\gamma)$ 1.214, $CH_2(\delta)$ 1.551; $CH_2(\varepsilon)$ 2.872; $NH_2(\phi)$ 7.480 |
| $Arg^7$ | 8.717 | 4.542 | 1.708, 1.763 | $CH_2(\gamma)$ 1.505, 1.622; $CH_2(\delta)$ 3.173; $NH(\varepsilon)$ 7.240 |
| D-$Arg^8$ | 9.208 | 4.118 | 1.821 | $CH_2(\gamma)$ 1.629, 1.713, $CH_2(\delta)$ 3.220 $NH(\varepsilon)$ 7.210 |
| $Gly^9$ | 8.987 | 3.944 | - | - |
| $Gln^{10}$ | 7.728 | 4.703 | 2.020, 2.117 | $CH_2(\gamma)$ 2.301, 2.356; $NH_2(\varepsilon)$ 7.590, 6.920 |
| $Thr^{11}$ | 8.941 | 4.625 | 3.960 | $CH_3(\gamma)$ 1.078 |
| $Arg^{12}$ | 8.740 | 4.716 | 1.770, 1.827 | $CH_2(\gamma)$ 1.554, 1.633; $CH_2(\delta)$ 3.181; $NH(\varepsilon)$ 7.200 |
| Cys13 | 9.076 | 5.591 | 2.607, 2.983 | - |
| $Ile^{14}$ | 8.858 | 4.516 | 1.828 | $CH_2(\gamma)$ 1.140, 1.422; $CH_3(\gamma^2)$ 0.895; $CH_3(\delta^1)$ 0.834 |
| $Arg^{15}$ | 8.586 | 4.846 | 1.675 1.814 | $CH_2(\gamma)$ 1.471, 1.555; $CH_2(\delta)$ 3.117, 3.194 |
| $Ile^{16}$ | 8.821 | 4.546 | 1.779 | $CH_2(\gamma)$ 1.068, 1.415 $CH_3(\gamma^2)$ 0.859; $CH_3(\delta^1)$ 0.815 |
| $Pro^{17}$ | - | 4.747 | 1.916, 2.286 | $CH_2(\gamma)$ 2.014, 2.130; $CH_2(\delta)$ 3.521, 3.862 |
| $Pro^{18}$ | - | 4.547 | 2.129, 2.235 | $CH_2(\gamma)$1.889, 2.106; $CH_2(\delta)$ 3.740, 3.975 |

**Table 73.** $^1$H-NMR assignment of **L4-07**. Chemical shifts obtained relative to the internal TSP standard at 298K and pH=2.3 in $H_2O/D_2O$ (9:1).

| **L4-07** | **NH** | **α** | **β** | **Others** |
|---|---|---|---|---|
| $Arg^1$ | 7.751 | 4.603 | 1.836 | $CH_2(\gamma)$ 1.600; $CH_2(\delta)$ 3.199; $NH(\varepsilon)$ 7.280 |
| $Val^2$ | 8.475 | 4.664 | 1.894 | $CH_3(\gamma^2)$ 0.837, |
| $Arg^3$ | 8.918 | 4.716 | 1.665, 1.812 | $CH_2(\gamma)$ 1.490; $CH_2(\delta)$ 3.156; $NH(\varepsilon)$ 7.230 |
| $Cys^4$ | 9.183 | 5.508 | 2.796, 2.985 | - |
| $Arg^5$ | 8.653 | 4.585 | 1.823, 1.901 | $CH_2(\gamma)$ 1.678; $CH_2(\delta)$ 3.134, 3.194; $NH(\varepsilon)$ 7.230 |
| $Arg^6$ | 8.621 | 4.354 | 1.802, 1.878 | $CH_2(\gamma)$ 1.613, 1.670; $CH_2(\delta)$ 3.151; $NH_2(\varepsilon)$ 7.480 |
| $Lys^7$ | 8.186 | 4.442 | 1.652, 1.734 | $CH_2(\gamma)$ 1.317, 1.354; $CH_2(\delta)$ 1.641; $CH_2(\varepsilon)$ 2.972 |
| Gly | 9.713 | 3.802, 4.307 | - | - |
| $Pro^9$ | - | 4.554 | 2.042, 2.260 | $CH_2(\gamma)$1.989; $CH_2(\delta)$ 3.794, 3.654 |
| $Gly^{10}$ | 8.136 | 3.882, 4.158 | - | - |
| $Gln^{11}$ | 8.330 | 4.581 | 1.883, 2.020 | $CH_3(\gamma)$ 2.229; $NH_2(\varepsilon)$ 7.525 |
| $Arg^{12}$ | 8.644 | 4.517 | 1.768, 1.886 | $CH_2(\gamma)$ 1.674; $CH_2(\delta)$ 3.181; $NH(\varepsilon)$ 7.200 |
| $Cys^{13}$ | 8.906 | 5.466 | 2.613, 3.045 | - |
| $Ile^{14}$ | 8.758 | 4.494 | 1.843 | $CH_2(\gamma)$ 1.150, 1.433; $CH_3(\gamma^2)$ 0.896; $CH_3(\delta^1)$ 0.849 |
| $Arg^{15}$ | 8.599 | 4.327 | 1.682 1.819 | $CH_2(\gamma)$ 1.471, 1.555; $CH_2(\delta)$ 3.117 |
| $Ile^{16}$ | 8.800 | 4.552 | 1.781 | $CH_2(\gamma)$ 1.077, 1.424, $CH_3(\gamma^2)$ 0.866; $CH_3(\delta^1)$ 0.824 |
| $Pro^{17}$ | - | 4.747 | 1.914, 2.286 | $CH_2(\gamma)$ 2.010, 2.129; $CH_2(\delta)$ 3.528, 3.865 |
| $Pro^{18}$ | - | 4.539 | 2.110, 2.232 | $CH_2(\gamma)$1.897, 2.104; $CH_2(\delta)$ 3.735, 3.978 |

**Table 74**. $^1$H-NMR assignment of **L4-09**. Chemical shifts obtained relative to the internal TSP standard at 298K and pH=2.3 in $H_2O/D_2O$ (9:1).

| **L4-09** | **NH** | **α** | **β** | **Others** |
|---|---|---|---|---|
| $Arg^1$ | 7.696 | 4.505 | 1.898 | $CH_2(\gamma)$ 1.557, 1.650; $CH_2(\delta)$ 3.222; NH(ε) 7.308 |
| $Cys^2$ | 8.962 | 5.499 | 2.584, 3.083 | - |
| $Arg^3$ | 8.769 | 4.629 | 1.773, 1.820 | $CH_2(\gamma)$ 1.524, 1.641; $CH_2(\delta)$ 3.156; NH(ε) 7.230 |
| $Val^4$ | 8.694 | 4.520 | 1.904 | $CH_3(\gamma^2)$ 0.810, 0.877 |
| $Arg^5$ | 8.952 | 4.629 | 1.804, 1.848 | $CH_2(\gamma)$ 1.549, 1.655; $CH_2(\delta)$ 3.144, 3.199; NH(ε) 7.334 |
| $Gln^6$ | 8.891 | 4.653 | 1.920, 1.967 | $CH_2(\gamma)$ 2.244; $NH_2(\varepsilon)$ 7.564 |
| $Arg^7$ | 8.804 | 4.497 | 1.731, 1.760 | $CH_2(\gamma)$ 1.494, 1.616; $CH_2(\delta)$ 3.165, 3.235; NH(ε) 7.370 |
| $Lys^8$ | 9.382 | 4.030 | 1.775, 1.958 | $CH_2(\gamma)$ 1.406; $CH_2(\delta)$ 1.686; $CH_2(\varepsilon)$ 2.975 |
| $Gly^9$ | 8.732 | 3.703, 4.112 | - | - |
| $Arg^{10}$ | 7.918 | 4.573 | 1.749, 1.825 | $CH_2(\gamma)$ 1.615; $CH_2(\delta)$ 3.165, 3.235; NH(ε) 7.370 |
| $Ala^{11}$ | 8.638 | 4.508 | 1.274 | - |
| $Arg^{12}$ | 8.748 | 4.540 | 1.727, 1.792 | $CH_2(\gamma)$ 1.607; $CH_2(\delta)$ 3.181; NH(ε) 7.200 |
| $Lys^{13}$ | 8.805 | 4.628 | 1.667, 1.775 | $CH_2(\gamma)$ 1.257, $CH_2(\delta)$ 1.630; $CH_2(\varepsilon)$ 2.895 |
| $Ile^{14}$ | 8.972 | 4.400 | 1.797 | $CH_2(\gamma)$ 1.194, 1.429; $CH_3(\gamma^2)$ 0.874; $CH_3(\delta^1)$ 0.803 |
| $Cys^{15}$ | 9.117 | 5.393 | 2.953, 3.099 | - |
| $Ile^{16}$ | 8.670 | 4.563 | 1.891 | $CH_2(\gamma)$ 1.142, 1.522; $CH_3(\gamma^2)$ 0.904; $CH_3(\delta^1)$ 0.873 |
| $Pro^{17}$ | - | 4.756 | 1.846, 2.313 | $CH_2(\gamma)$ 2.004, 2.106; $CH_2(\delta)$ 3.609, 3.902 |
| $Pro^{18}$ | - | 4.519 | 2.134, 2.238 | $CH_2(\gamma)$1.906, 2.095; $CH_2(\delta)$ 3.727, 3.928 |

**Table 75.** $^1$H-NMR assignment of **L5-11**. Chemical shifts obtained relative to the internal TSP standard at 298K and pH=2.3 in $H_2O/D_2O$ (9:1).

| **L5-11** | **NH** | **α** | **β** | **Others** |
|---|---|---|---|---|
| $Arg^1$ | 7.755 | 4.611 | 1.835 | $CH_2(\gamma)$ 1.603; $CH_2(\delta)$ 3.203; NH(ε) 7.270 |
| $Val^2$ | 8.473 | 4.708 | 1.883 | $CH_3(\gamma^2)$ 0.839, |
| $Arg^3$ | 8.897 | 4.732 | 1.656, 1.807 | $CH_2(\gamma)$ 1.489; $CH_2(\delta)$ 3.164; NH(ε) 7.227 |
| $Cys^4$ | 9.155 | 5.659 | 2.628, 3.047 | - |
| $Arg^5$ | 8.855 | 4.633 | 1.804, 1.851 | $CH_2(\gamma)$ 1.528, 1.659; $CH_2(\delta)$ 3.135, 3.198; NH(ε) 7.190 |
| $Gln^6$ | 8.727 | 4.72 | 1.933 | $CH_2(\gamma)$ 2.190, 2.212 |
| $Arg^7$ | 8.888 | 4.527 | 1.672, 1.779 | $CH_2(\gamma)$ 1.472, 1.618; $CH_2(\delta)$ 3.160; NH(ε) 7.250 |
| $Lys^8$ | 9.379 | 3.949 | 1.804, 1.993 | $CH_2(\gamma)$ 1.405; $CH_2(\delta)$ 1.690; $CH_2(\varepsilon)$ 2.989 |
| $Gly^9$ | 8.623 | 3.669, 4.118 | - | - |
| $Arg^{10}$ | 7.793 | 4.635 | 1.766, 1.858 | $CH_2(\gamma)$ 1.597; $CH_2(\delta)$ 3.214; NH(ε) 7.189 |
| $Arg^{11}$ | 8.651 | 4.513 | 1.700, 1.760 | $CH_3(\gamma)$ 1.477 $CH_2(\delta)$ 3.146; NH(ε) 7.206 |
| $Ala^{12}$ | 8.842 | 4.722 | 1.361 | - |
| Cys13 | 8.891 | 5.555 | 2.634, 2.995 | - |
| $Ile^{14}$ | 8.863 | 4.523 | 1.813 | $CH_2(\gamma)$ 1.143, 1.435; $CH_3(\gamma^2)$ 0.904; $CH_3(\delta^1)$ 0.845 |
| $Arg^{15}$ | 8.563 | 4.864 | 1.679 1.816 | $CH_2(\gamma)$ 1.472, 1.563; $CH_2(\delta)$ 3.129, 3.201, NH(ε) 7.255 |
| $Ile^{16}$ | 8.816 | 4.550 | 1.784 | $CH_2(\gamma)$ 1.070, 1.416; $CH_3(\gamma^2)$ 0.867; $CH_3(\delta^1)$ 0.813 |
| $Pro^{17}$ | - | 4.762 | 2.290, 2.013 | $CH_2(\gamma)$ 2.013, 2.129; $CH_2(\delta)$ 3.523, 3.863 |
| $Pro^{18}$ | - | 4.550 | 2.111, 2.240 | $CH_2(\gamma)$1.895, 2.108; $CH_2(\delta)$ 3.741, 3.975 |

**Table 76.** $^1$H-NMR assignment of **R1-06**. Chemical shifts obtained relative to the internal TSP standard at 298K and pH=2.3 in $H_2O/D_2O$ (9:1).

| **R1-06** | **NH** | **α** | **β** | **Others** |
|---|---|---|---|---|
| Trp[1] | 8.044 | 4.469 | 3.311 | $CH_2$(δ) 7.291; CH(ε3) 7.587; CH(ε1) 10.160, CH(ζ3) 7.119; CH(ζ2) 7.480; CH(η2) 7.221 |
| Arg[2] | 7.772 | 4.028 | 1.537, 1.685 | $CH_2$(γ) 1.302; $CH_2$(δ) 3.086; NH(ε) 7.153 |
| Arg[3] | 7.895 | 4.209 | 1.695 1.882 | $CH_2$(γ) 1.541, 1.581; $CH_2$(δ) 3.173; NH(ε) 7.2 |
| Arg[4] | 7.903 | 4.328 | 1.575, 1.876 | $CH_2$(γ) 1.598; $CH_2$(δ) 3.115; NH(ε) 7.14 |
| Ala[5] | 8.249 | 4.370 | 1.389 | - |
| Gly[6] | 8.380 | 3.886, 4.126 | - | - |
| Lys[7] | 8.439 | 4.283 | 1.744 1.860 | $CH_2$(γ) 1.393, 1.450; $CH_2$(δ) 1.669; $CH_2$(ε) 2.990 |
| Gln[8] | 8.370 | 4.297 | 1.966, 2.115 | $CH_2$(γ) 2.358, 1.687; $CH_2$(ε) 7.484, 6.870 |
| Arg[9] | 8.392 | 4.221 | 1.718 1.789 | $CH_2$(γ) 1.569; $CH_2$(δ) 3.147; NH(ε) 7.16 |
| Asn[10] | 8.215 | 4.641 | 2.716, 2.793 | $CH_2$(γ) 7.525, 6.870 |
| Arg[11] | 8.224 | 4.168 | 1.618 | $CH_2$(γ) 1.243, 1.450; $CH_2$(δ) 2.947; NH(ε) 7.003 |
| Arg[12] | 8.181 | 4.455 | 1.747, 1.867 | $CH_2$(γ)1.611, 1.651; $CH_2$(δ) 3.202; NH(ε) 7.19 |
| Pro[13] | - | 4.718 | 1.947, 2.337 | $CH_2$(γ) 2.080, 2.127; $CH_2$(δ) 3.747 |
| Pro[14] | - | 4.441 | 1.753, 2.092 | $CH_2$(γ)1.605, 1.925; $CH_2$(δ) 3.636, 3.808 |

**Table 77.** $^1$H-NMR assignment of **R2-14**. Chemical shifts obtained relative to the internal TSP standard at 298K and pH=2.3 in $H_2O/D_2O$ (9:1).

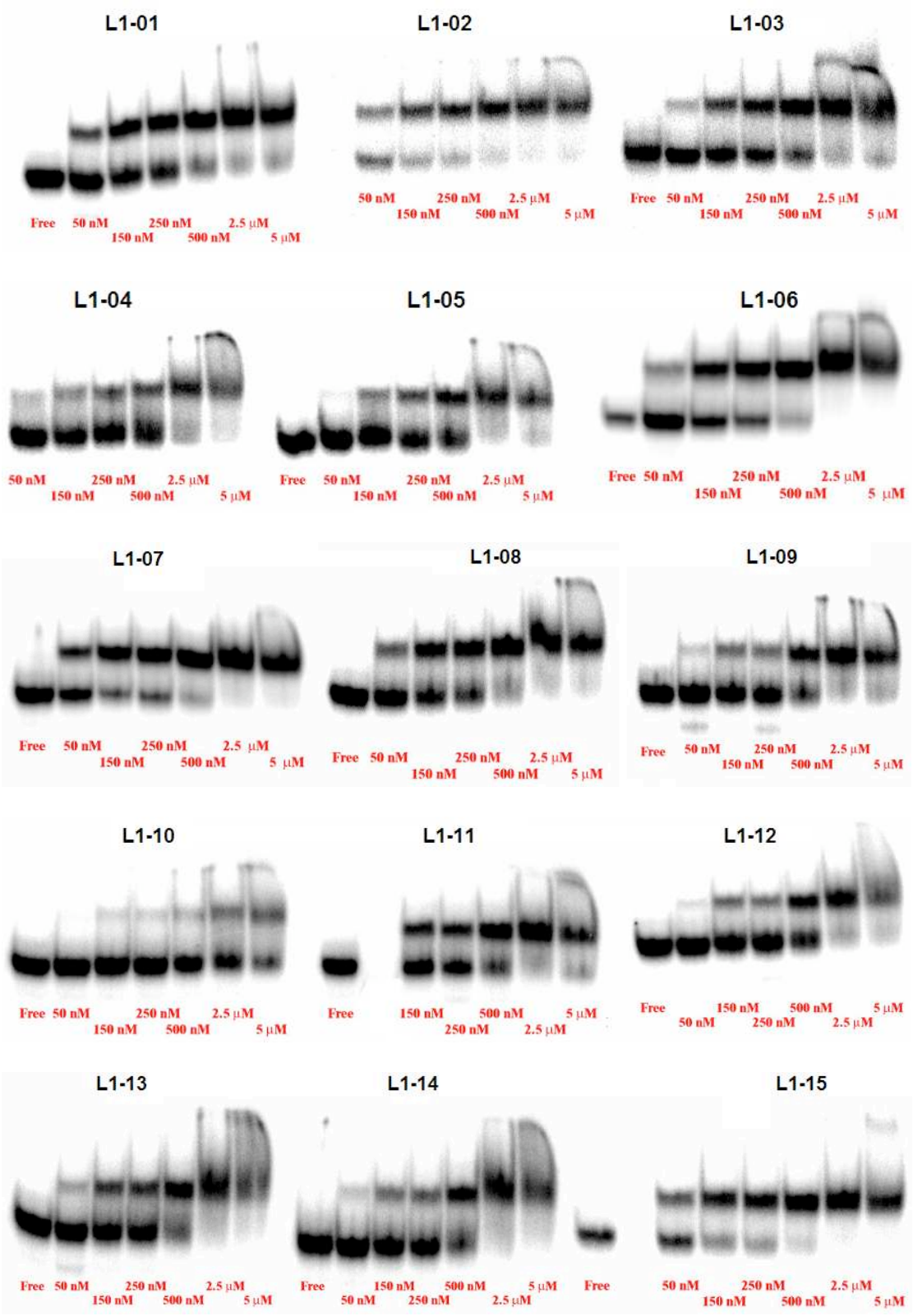
| **R2-14** | **NH** | **α** | **β** | **Others** |
|---|---|---|---|---|
| Arg[1] | 7.693 | 4.450 | 1.911 | $CH_2$(γ) 1.605; 1.667; $CH_2$(δ) 3.234; NH(ε) 7.260 |
| Cys[2] | 8.584 | 5.247 | 2.737, 2.994 | - |
| Arg[3] | 8.638 | 4.532 | 1.818, | $CH_2$(γ) 1.562, 1.624; $CH_2$(δ) 3.208; NH(ε) 7.200 |
| Arg[4] | 8.577 | 4.699 | 1.685, 1.773 | $CH_2$(γ) 1.445, 1.536; $CH_2$(δ) 3.132; NH(ε) 7.217 |
| Ala[5] | 8.423 | 4.515 | 1.340 | - |
| Lys[6] | 9.004 | 4.056 | 1.825, 1.926 | $CH_2$(γ) 1.420; $CH_2$(δ) 1.694; $CH_2$(ε) 3.001 |
| Gly[7] | 8.615 | 3.720, 4.119 | - | - |
| Gln[8] | 7.946 | 4.529 | 1.965, 2.108 | $CH_2$(γ) 2.324; $CH_2$(ε) 7.572, 6.905 |
| Arg[9] | 8.693 | 4.514 | 1.662, 1.834 | $CH_2$(γ) 1.508; $CH_2$(δ) 3.163; NH(ε) 7.205 |
| Arg[10] | 8.677 | 4.582 | 1.759, 1.840 | $CH_2$(γ) 1.583, 1.637; $CH_2$(δ) 3.149, 3.187; NH(ε) 7.174 |
| Cys[11] | 8.786 | 5.215 | 2.940, 3.128 | - |
| Arg[12] | 8.592 | 4.869 | 1.747, 1.839 | $CH_2$(γ) 1.578, 1.664; $CH_2$(δ) 3.223; NH(ε) 7.2 |
| Pro[13] | - | 4.722 | 1.890, 2.303 | $CH_2$(γ) 2.015, 2.097; $CH_2$(δ) 3.618, 3.812 |
| Pro[14] | - | 4.483 | 2.089, 2.274 | $CH_2$(γ)1.931, 2.090; $CH_2$(δ) 3.719, 3.925 |

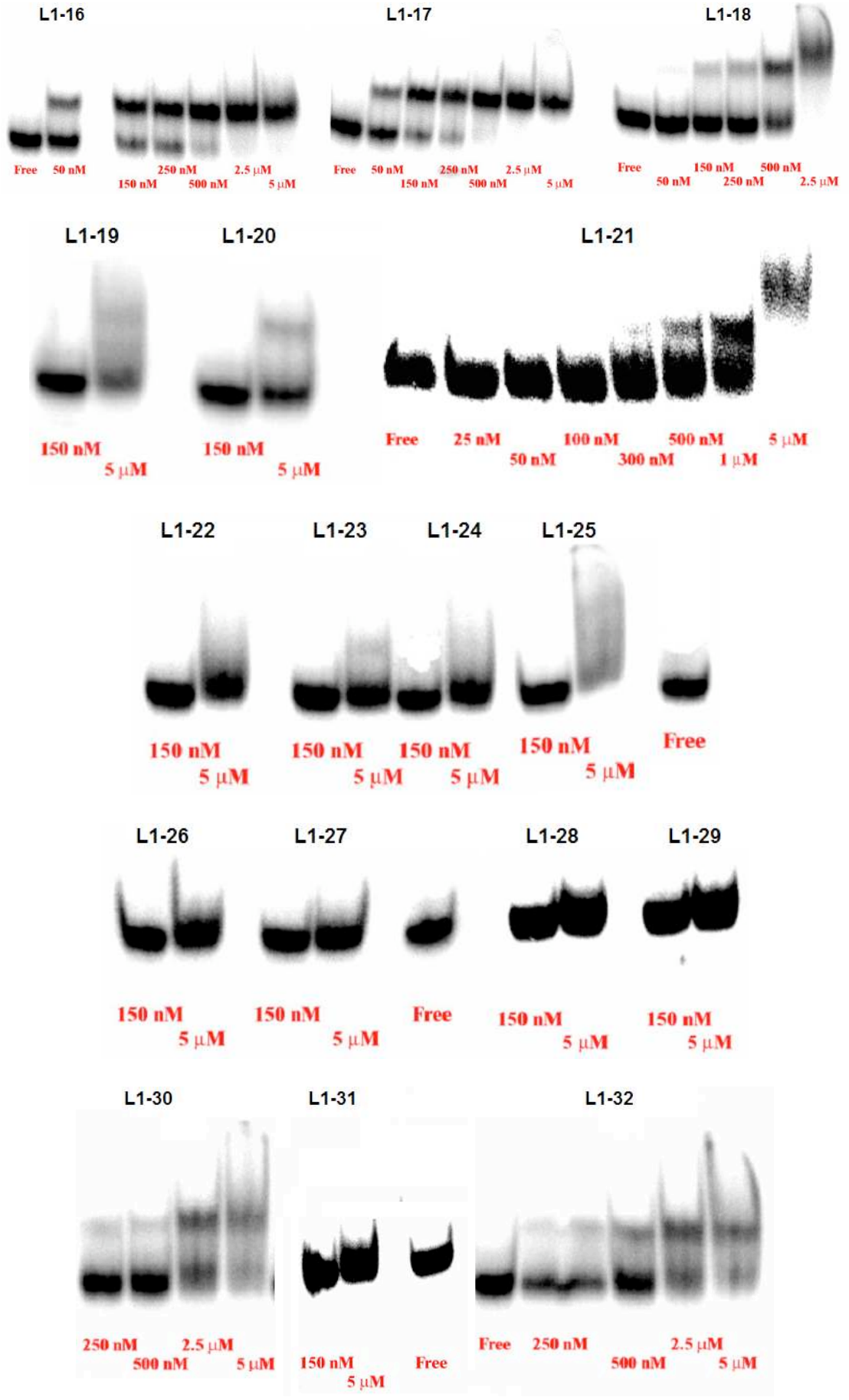
# Appendix 7: EMSA images for BIV/HIV Tat and Rev peptidomimetics

**Table 78.** Library **L1** of BIV Tat mimetics. Residues 1 and 12 are attached to the D-Pro-L-Pro template.

| **Mimetic** | **Position** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** |
| **BIV2** | **R** | **V** | **R** | **T** | **R** | **G** | **K** | **R** | **R** | **I** | **R** | **V** |
| **L1-01** | R | V | R | T | R | G | K | R | R | I | R | **L** |
| **L1-02** | R | V | R | T | R | G | K | R | R | I | R | **I** |
| **L1-03** | R | V | R | T | R | G | K | R | R | I | R | **F** |
| **L1-04** | R | V | R | T | R | G | K | R | R | I | R | **T** |
| **L1-05** | R | V | R | T | R | G | K | R | R | I | R | **N** |
| **L1-06** | R | V | R | T | R | G | K | R | R | I | R | **R** |
| **L1-07** | R | **T** | R | T | R | G | K | R | R | I | R | V |
| **L1-08** | R | **N** | R | T | R | G | K | R | R | I | R | V |
| **L1-09** | R | V | R | T | R | G | K | R | R | **L** | R | V |
| **L1-10** | R | V | R | T | R | G | K | R | R | **Cha** | R | V |
| **L1-11** | R | V | R | T | R | G | K | R | R | **F** | R | V |
| **L1-12** | R | V | R | T | R | G | K | R | R | **Y** | R | V |
| **L1-13** | R | V | R | T | R | G | K | R | R | **N** | R | V |
| **L1-14** | R | V | R | T | R | G | K | R | R | **Q** | R | V |
| **L1-15** | R | V | R | **Q** | R | G | K | R | R | I | R | V |
| **L1-16** | R | V | R | **V** | R | G | K | R | R | I | R | V |
| **L1-17** | R | V | R | **Y** | R | G | K | R | R | I | R | V |
| **L1-18** | R | V | R | **K** | R | G | K | R | R | I | R | V |
| **L1-19** | R | V | **K** | T | R | G | K | R | R | I | R | V |
| **L1-20** | R | V | **U** | T | R | G | K | R | R | I | R | V |
| **L1-21** | R | V | **O** | T | R | G | K | R | R | I | R | V |
| **L1-22** | R | V | R | T | **K** | G | K | R | R | I | R | V |
| **L1-23** | R | V | R | T | **U** | G | K | R | R | I | R | V |
| **L1-24** | R | V | R | T | **O** | G | K | R | R | I | R | V |
| **L1-25** | R | V | R | T | **N** | G | K | R | R | I | R | V |
| **L1-26** | R | V | R | T | **Q** | G | K | R | R | I | R | V |
| **L1-27** | R | V | R | T | **I** | G | K | R | R | I | R | V |
| **L1-28** | R | V | R | T | **L** | G | K | R | R | I | R | V |
| **L1-29** | R | V | R | T | **Y** | G | K | R | R | I | R | V |
| **L1-30** | R | V | R | T | R | G | K | **K** | R | I | R | V |
| **L1-31** | R | V | R | T | R | G | K | **U** | R | I | R | V |
| **L1-32** | R | V | R | T | R | G | K | **O** | R | I | R | V |
| **L1-33** | R | V | R | T | R | G | K | **N** | R | I | R | V |
| **L1-34** | R | V | R | T | R | G | K | **F** | R | I | R | V |
| **L1-35** | R | V | R | T | R | G | K | **Y** | R | I | R | V |
| **L1-36** | R | V | R | T | R | G | K | R | **K** | I | R | V |
| **L1-37** | R | V | R | T | R | G | K | R | **U** | I | R | V |
| **L1-38** | R | V | R | T | R | G | K | R | **O** | I | R | V |
| **L1-39** | R | V | R | T | R | G | K | R | **N** | I | R | V |
| **L1-40** | R | V | R | T | R | G | K | R | **Q** | I | R | V |
| **L1-41** | R | V | R | T | R | G | K | R | R | I | **K** | V |
| **L1-42** | R | V | R | T | R | G | K | R | R | I | **U** | V |
| **L1-43** | R | V | R | T | R | G | K | R | R | I | **O** | V |
| **L1-44** | R | V | R | T | R | G | K | R | R | I | **N** | V |
| **L1-45** | R | V | R | T | R | G | K | R | R | I | **Q** | V |

**Figure 80.** Gel band shift assay of mimetics from library **L1**. Peptides and **BIV TAR RNA** (~2 nM) were incubated at the peptide concentrations shown.




L1-16
Free
50 nM
150 nM
250 nM
500 nM
2.5 μM
5 μM
L1-17
Free
50 nM
150 nM
250 nM
500 nM
2.5 μM
5 μM
L1-18
Free
50 nM
150 nM
250 nM
500 nM
2.5 μM
L1-19
150 nM
5 μM
L1-20
150 nM
5 μM
L1-21
Free
25 nM
50 nM
100 nM
300 nM
500 nM
1 μM
5 μM
L1-22
150 nM
5 μM
L1-23
150 nM
5 μM
L1-24
150 nM
5 μM
L1-25
150 nM
5 μM
Free
L1-26
150 nM
5 μM
L1-27
150 nM
5 μM
Free
L1-28
150 nM
5 μM
L1-29
150 nM
5 μM
L1-30
250 nM
500 nM
2.5 μM
5 μM
L1-31
150 nM
5 μM
Free
L1-32
Free
250 nM
500 nM
2.5 μM
5 μM

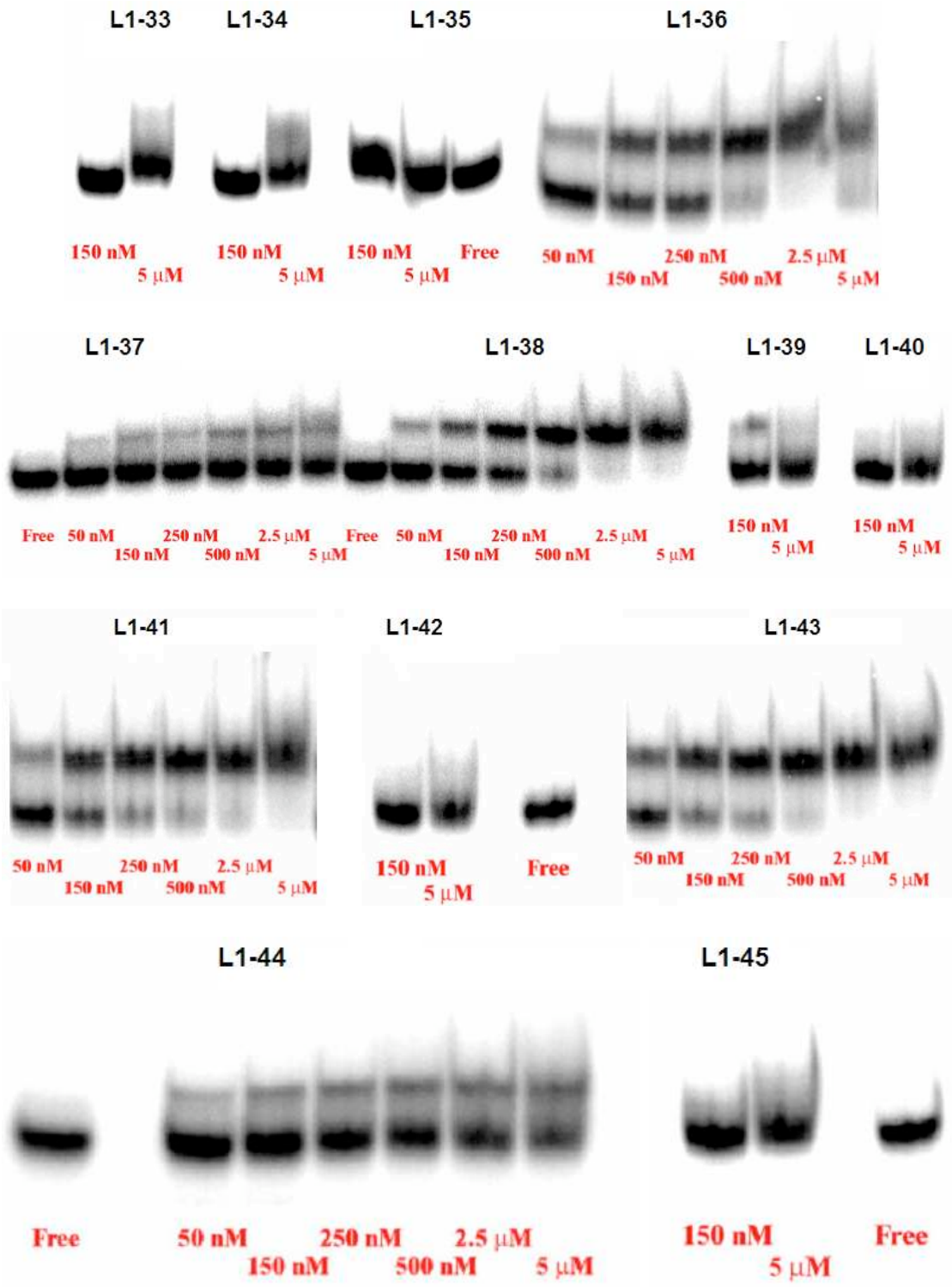
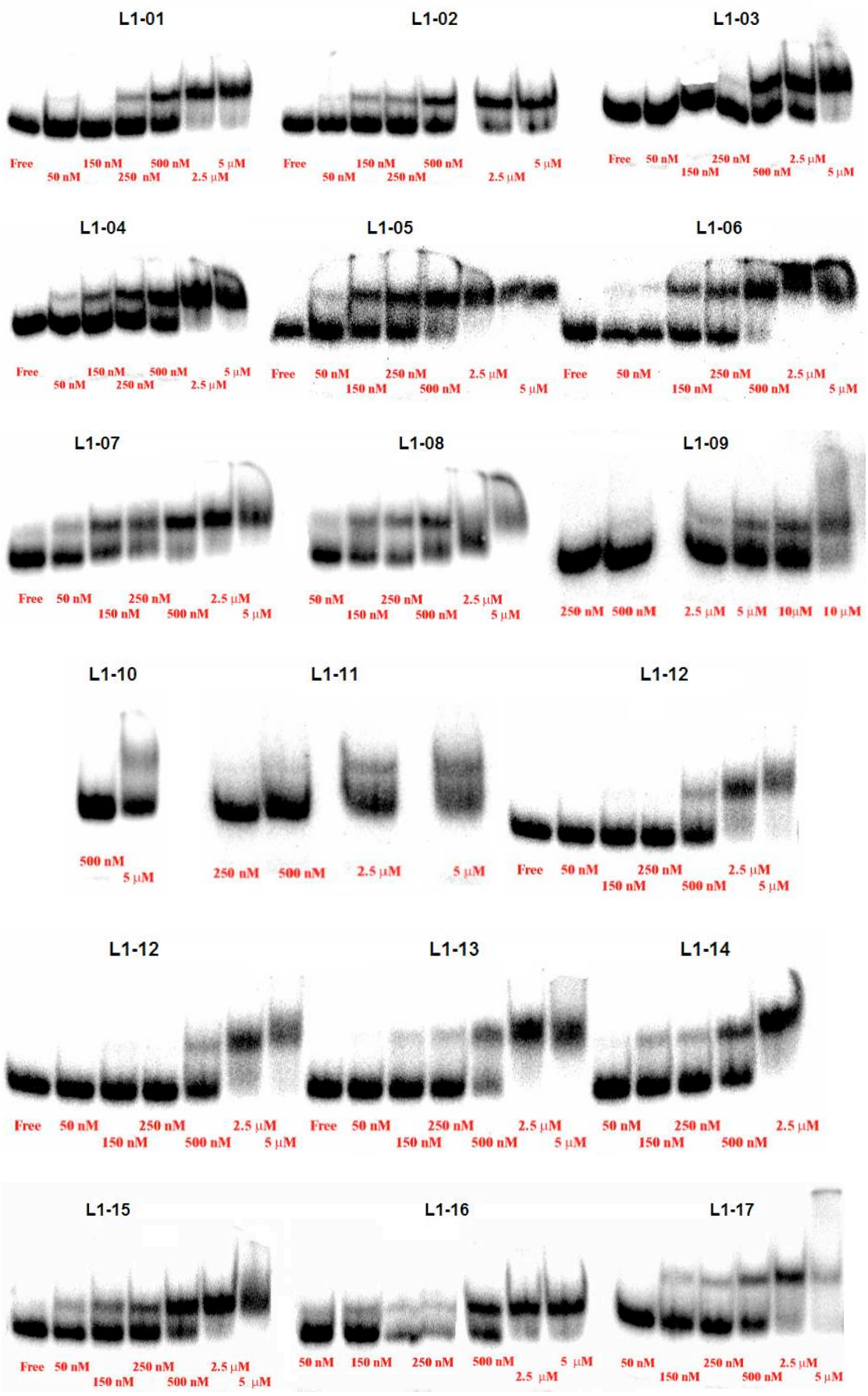
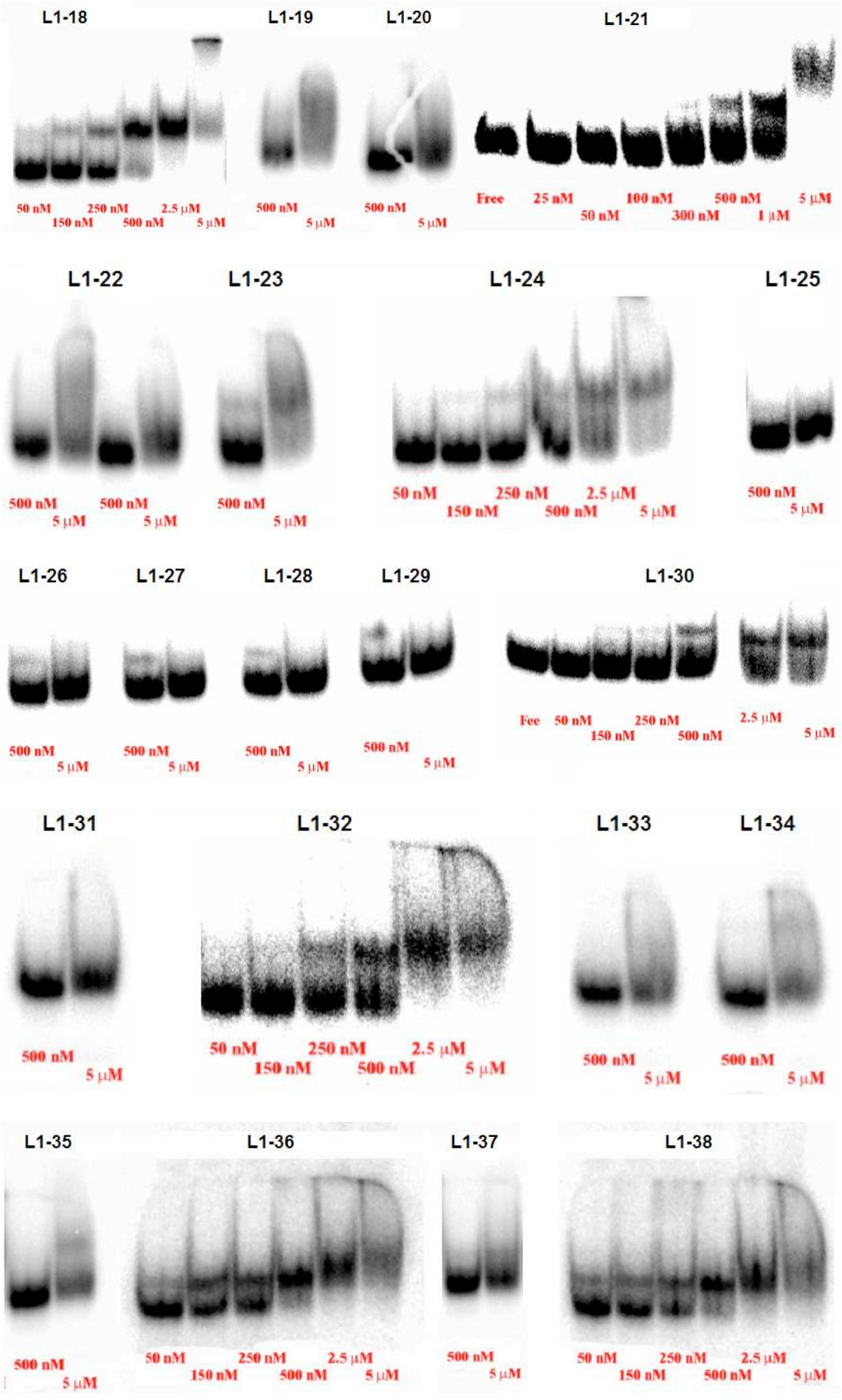
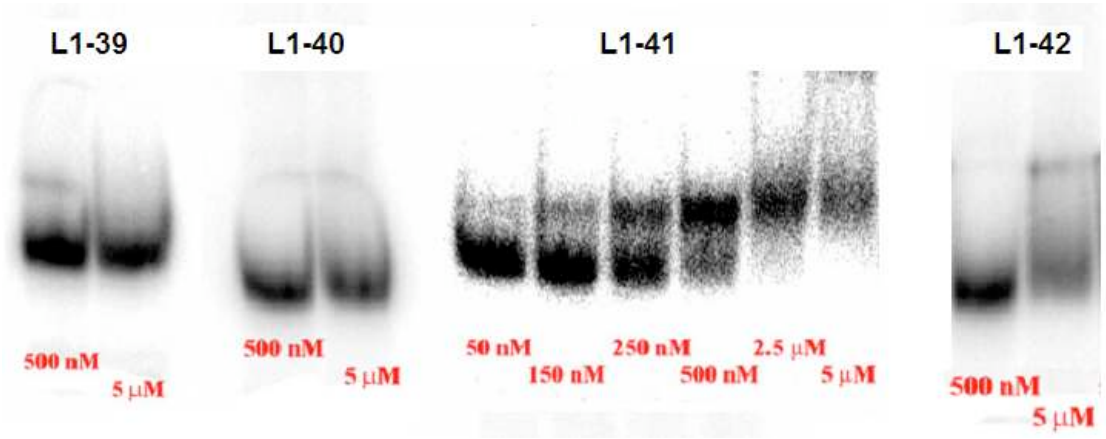

L1-33
L1-34
L1-35
L1-36
150 nM
5 µM
150 nM
5 µM
150 nM
5 µM
Free
50 nM
150 nM
250 nM
500 nM
2.5 µM
5 µM
L1-37
L1-38
L1-39
L1-40
Free
50 nM
150 nM
250 nM
500 nM
2.5 µM
5 µM
Free
50 nM
150 nM
250 nM
500 nM
2.5 µM
5 µM
150 nM
5 µM
150 nM
5 µM
L1-41
L1-42
L1-43
50 nM
150 nM
250 nM
500 nM
2.5 µM
5 µM
150 nM
5 µM
Free
50 nM
150 nM
250 nM
500 nM
2.5 µM
5 µM
L1-44
L1-45
Free
50 nM
150 nM
250 nM
500 nM
2.5 µM
5 µM
150 nM
5 µM
Free

**Figure 81.** Gel band shift assay of mimetics from library **L1**. Peptides and **HIV TAR RNA** (~2 nM) were incubated at the peptide concentrations shown.




L1-18
50 nM
150 nM
250 nM
500 nM
2.5 µM
5 µM
L1-19
500 nM
5 µM
L1-20
500 nM
5 µM
L1-21
Free
50 nM
25 nM
100 nM
300 nM
500 nM
1 µM
5 µM
L1-22
500 nM
5 µM
500 nM
5 µM
L1-23
500 nM
5 µM
L1-24
50 nM
150 nM
250 nM
500 nM
2.5 µM
5 µM
L1-25
500 nM
5 µM
L1-26
500 nM
5 µM
L1-27
500 nM
5 µM
L1-28
500 nM
5 µM
L1-29
500 nM
5 µM
L1-30
Fee
50 nM
150 nM
250 nM
500 nM
2.5 µM
5 µM
L1-31
500 nM
5 µM
L1-32
50 nM
150 nM
250 nM
500 nM
2.5 µM
5 µM
L1-33
500 nM
5 µM
L1-34
500 nM
5 µM
L1-35
500 nM
5 µM
L1-36
50 nM
150 nM
250 nM
500 nM
2.5 µM
5 µM
L1-37
500 nM
5 µM
L1-38
50 nM
150 nM
250 nM
500 nM
2.5 µM
5 µM


L1-39
500 nM
5 μM
L1-40
500 nM
5 μM
L1-41
50 nM
150 nM
250 nM
500 nM
2.5 μM
5 μM
L1-42
500 nM
5 μM

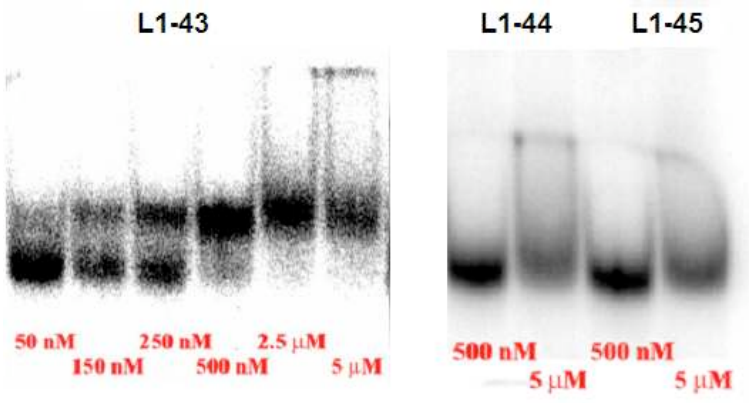
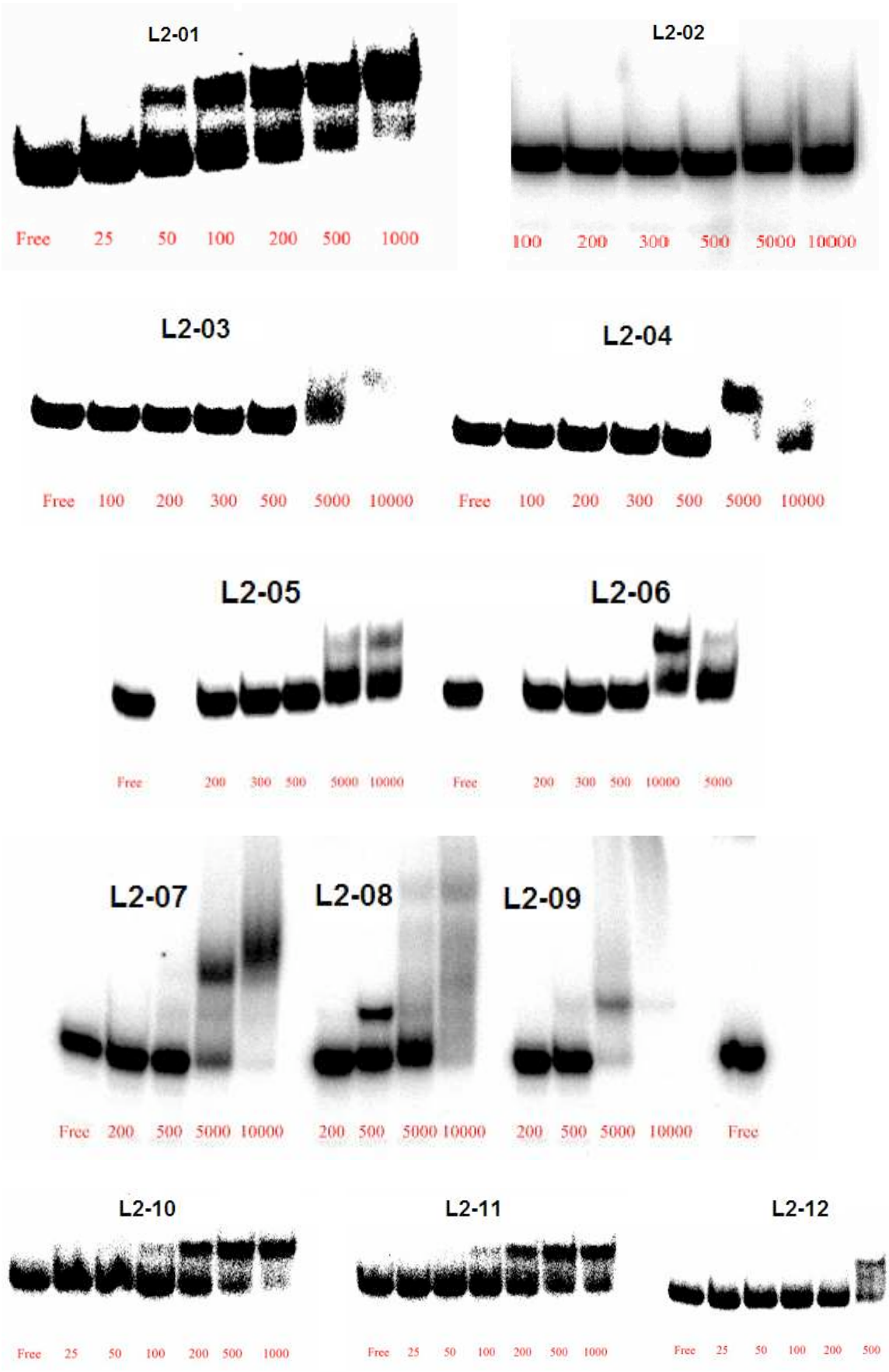
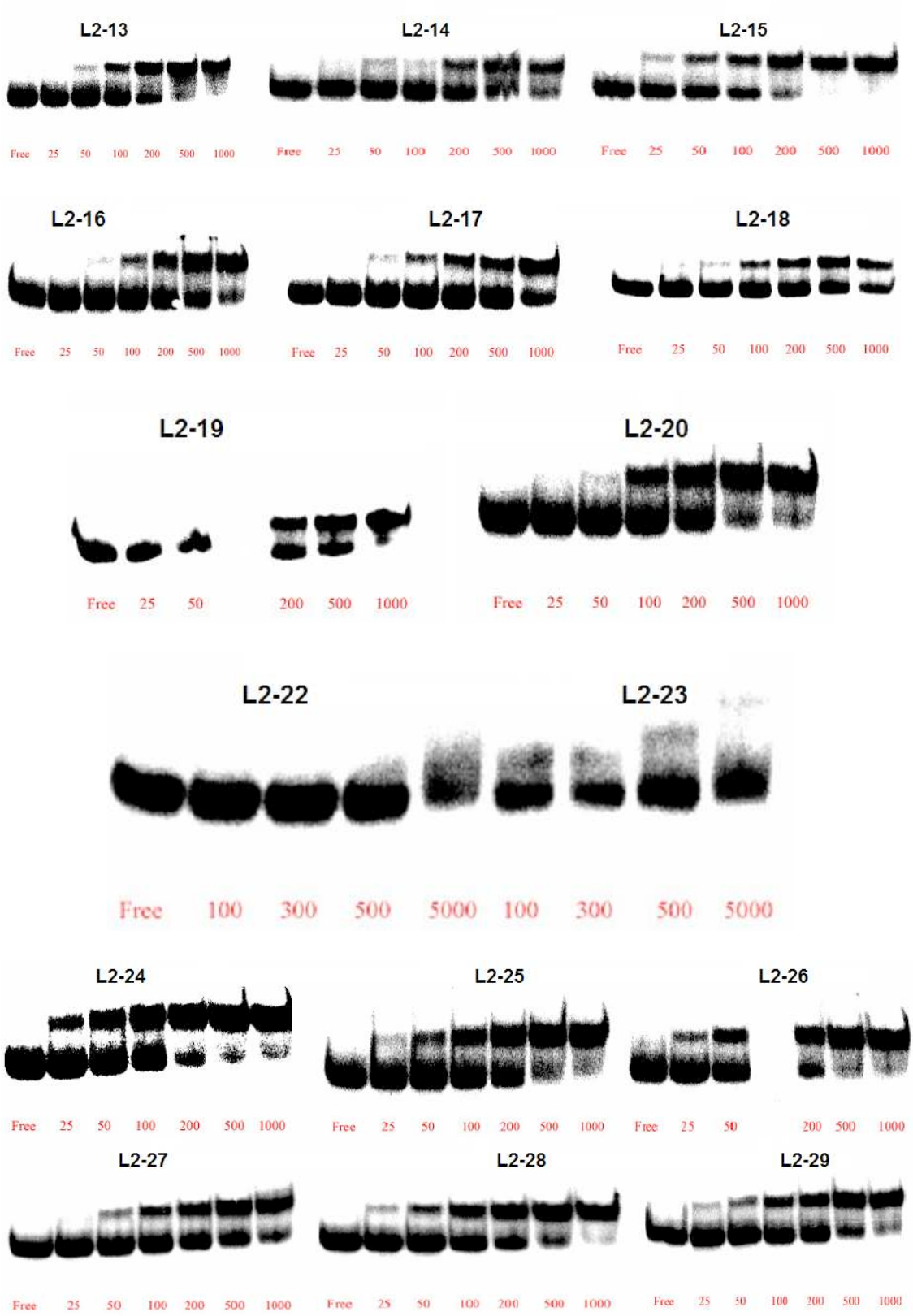

L1-43
50 nM
150 nM
250 nM
500 nM
2.5 μM
5 μM
L1-44
500 nM
5 μM
L1-45
500 nM
5 μM

**Table 79.** Library **L2** of BIV Tat mimetics. Residues 1 and 12 are attached to the D-Pro-L-Pro template.

| **Mimetic** | **Position** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** |
| L2-01 | R | V | R | T | R | G | K | R | R | I | R | **I** |
| L2-02 | **U** | V | R | T | R | G | K | R | R | I | R | **I** |
| L2-03 | **K** | V | R | T | R | G | K | R | R | I | R | **I** |
| L2-04 | **O** | V | R | T | R | G | K | R | R | I | R | **I** |
| L2-05 | **N** | V | R | T | R | G | K | R | R | I | R | **I** |
| L2-06 | **Q** | V | R | T | R | G | K | R | R | I | R | **I** |
| L2-07 | **Y** | V | R | T | R | G | K | R | R | I | R | **I** |
| L2-08 | **L** | V | R | T | R | G | K | R | R | I | R | **I** |
| L2-09 | **W** | V | R | T | R | G | K | R | R | I | R | **I** |
| L2-10 | R | **Q** | R | T | R | G | K | R | R | I | R | **I** |
| L2-11 | R | **Y** | R | T | R | G | K | R | R | I | R | **I** |
| L2-12 | R | **W** | R | T | R | G | K | R | R | I | R | **I** |
| L2-13 | R | **L** | R | T | R | G | K | R | R | I | R | **I** |
| L2-14 | R | V | R | T | R | **N** | **G** | R | R | I | R | **I** |
| L2-15 | R | V | R | T | R | **K** | **G** | R | R | I | R | **I** |
| L2-16 | R | V | R | T | R | G | **R** | R | R | I | R | **I** |
| L2-17 | R | V | R | T | R | G | **Y** | R | R | I | R | **I** |
| L2-18 | R | V | R | T | R | G | **Q** | R | R | I | R | **I** |
| L2-19 | R | V | R | T | R | G | **N** | R | R | I | R | **I** |
| L2-20 | R | V | R | T | R | G | **O** | R | R | I | R | **I** |
| L2-21 | R | V | R | T | R | G | **U** | R | R | I | R | **I** |
| L2-22 | R | V | **Q** | T | R | G | K | R | R | I | R | **I** |
| L2-23 | R | V | **N** | T | R | G | K | R | R | I | R | **I** |
| L2-24 | R | **T** | R | T | R | G | K | R | R | I | R | **I** |
| L2-25 | R | **N** | R | T | R | G | K | R | R | I | R | **I** |
| L2-26 | R | V | R | T | R | G | K | R | R | I | **O** | **I** |
| L2-27 | R | V | R | T | R | G | K | R | R | I | **K** | **I** |
| L2-28 | R | V | R | T | R | G | K | R | **K** | I | **O** | **I** |
| L2-29 | R | V | R | **Q** | R | G | K | R | R | I | R | **I** |
| L2-30 | R | V | R | **V** | R | G | K | R | R | I | R | **I** |
| L2-31 | R | V | R | **Y** | R | G | K | R | R | I | R | **I** |
| L2-32 | R | **T** | R | **Q** | R | G | K | R | R | I | R | **I** |
| L2-33 | R | **T** | R | T | R | G | K | R | R | I | **O** | **I** |
| L2-34 | R | **T** | R | T | R | G | K | R | **K** | I | **O** | **I** |
| L2-35 | R | **T** | R | **Q** | R | G | K | R | R | I | **O** | **I** |
| L2-36 | R | **T** | R | **V** | R | G | K | R | R | I | **O** | **I** |
| L2-37 | R | **T** | R | **Y** | R | G | K | R | R | I | **O** | **I** |
| L2-38 | R | V | R | T | R | G | K | R | R | **F** | R | **I** |
| L2-39 | R | **T** | R | T | R | G | K | R | R | **F** | R | **I** |
| L2-40 | R | **T** | R | T | R | G | K | R | R | **F** | **O** | **I** |
| L2-41 | R | V | R | **Q** | R | G | K | R | R | **F** | **O** | **I** |
| L2-42 | R | **T** | R | **V** | R | G | K | R | **K** | **F** | **O** | **I** |
| L2-43 | R | **T** | R | **Y** | R | G | K | R | **K** | **F** | **O** | **I** |

**Figure 82.** Gel band shift assay of mimetics from library **L2**. Peptides and **BIV TAR RNA** (~2 nM) were incubated at the peptide concentrations shown.




L2-13
Free 25 50 100 200 500 1000
L2-14
Free 25 50 100 200 500 1000
L2-15
Free 25 50 100 200 500 1000
L2-16
Free 25 50 100 200 500 1000
L2-17
Free 25 50 100 200 500 1000
L2-18
Free 25 50 100 200 500 1000
L2-19
Free 25 50 200 500 1000
L2-20
Free 25 50 100 200 500 1000
L2-22
L2-23
Free 100 300 500 5000 100 300 500 5000
L2-24
Free 25 50 100 200 500 1000
L2-25
Free 25 50 100 200 500 1000
L2-26
Free 25 50 200 500 1000
L2-27
Free 25 50 100 200 500 1000
L2-28
Free 25 50 100 200 500 1000
L2-29
Free 25 50 100 200 500 1000

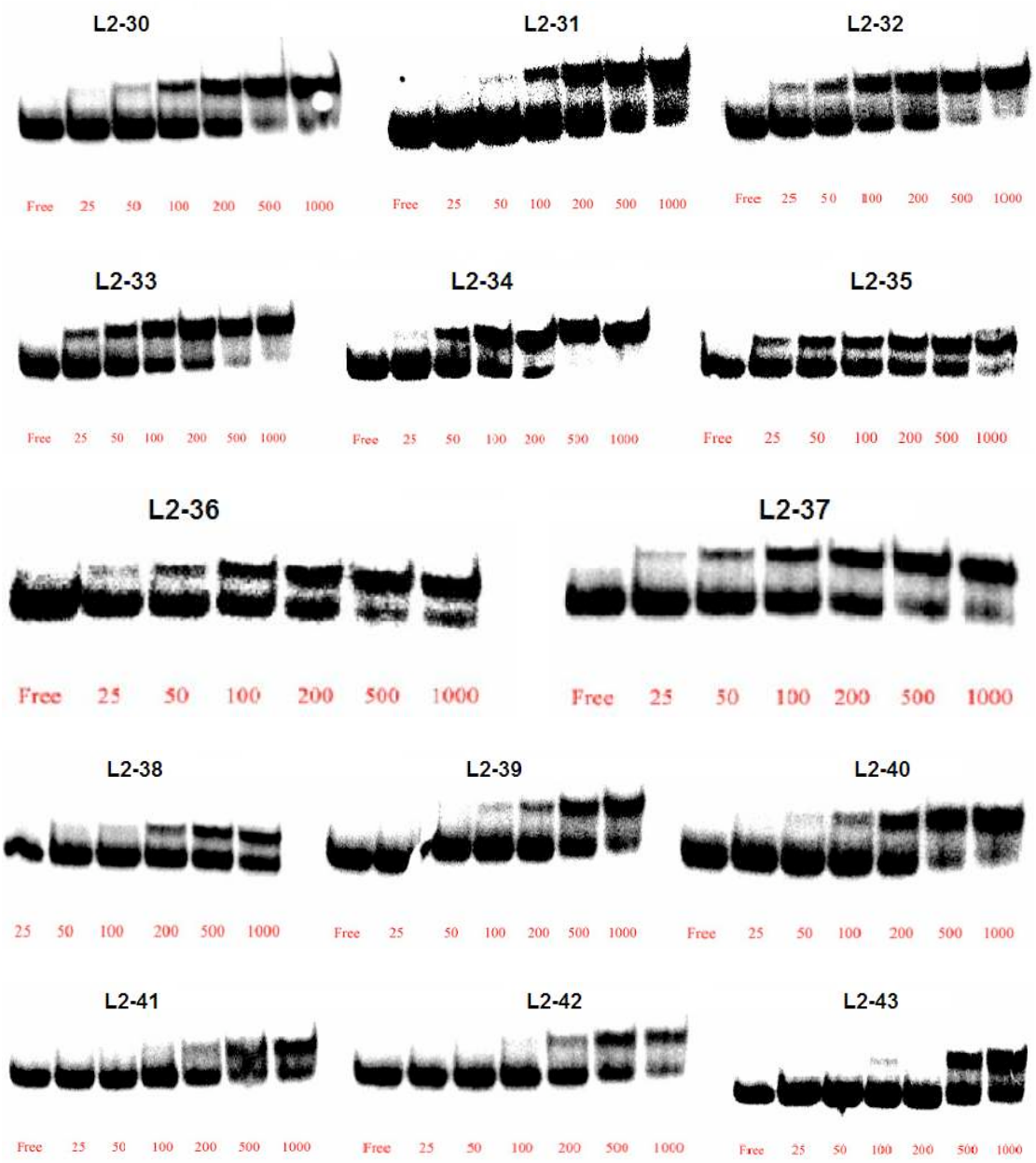
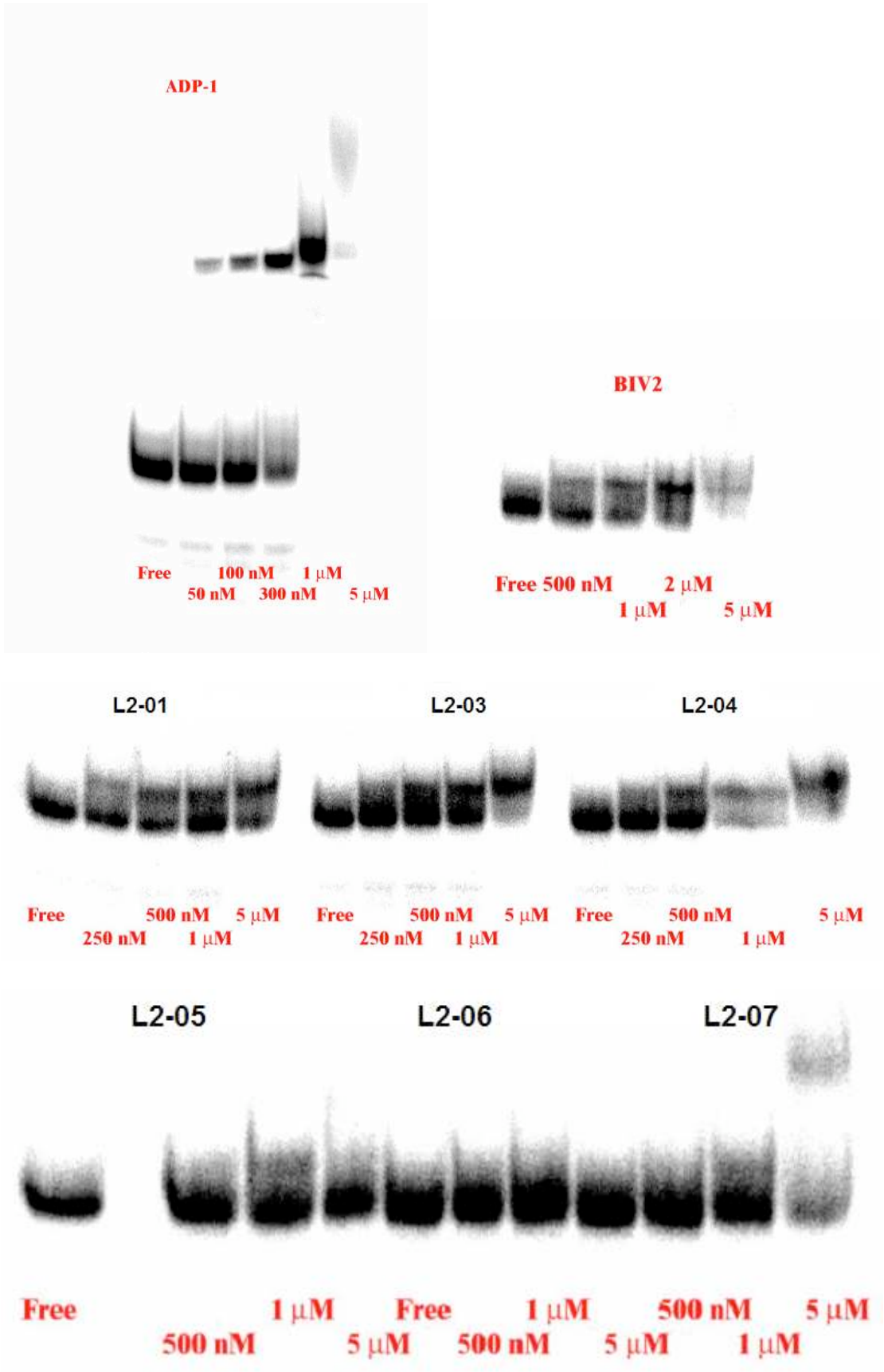
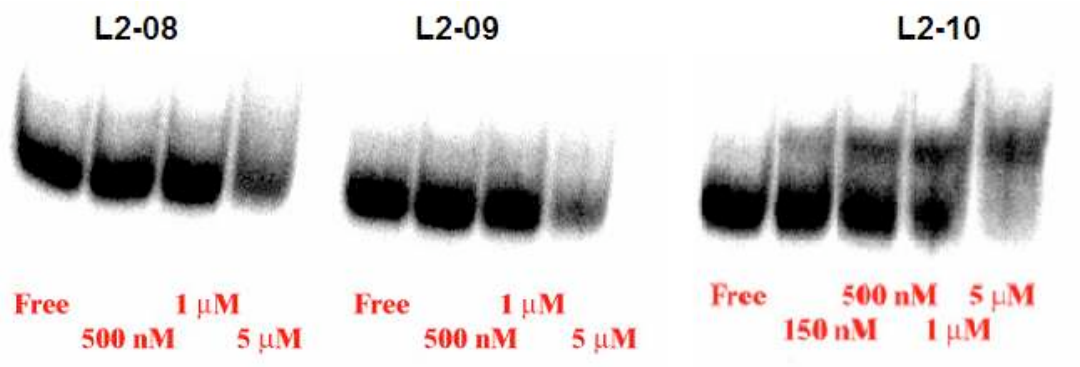
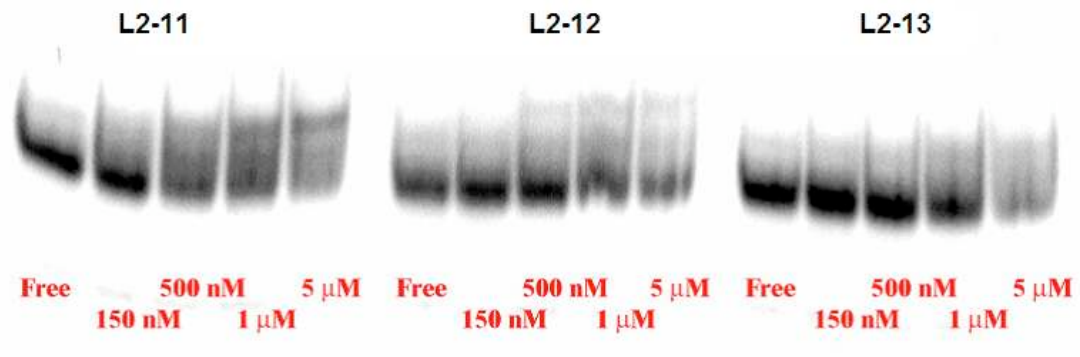
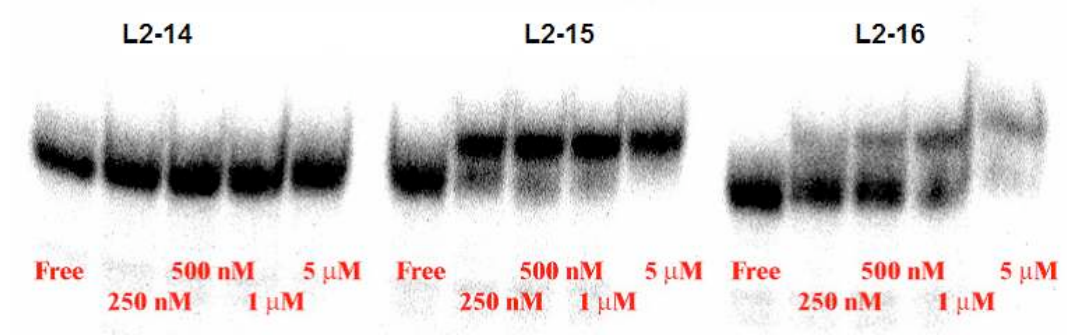

L2-30
Free 25 50 100 200 500 1000
L2-31
Free 25 50 100 200 500 1000
L2-32
Free 25 50 100 200 500 1000
L2-33
Free 25 50 100 200 500 1000
L2-34
Free 25 50 100 200 500 1000
L2-35
Free 25 50 100 200 500 1000
L2-36
Free 25 50 100 200 500 1000
L2-37
Free 25 50 100 200 500 1000
L2-38
25 50 100 200 500 1000
L2-39
Free 25 50 100 200 500 1000
L2-40
Free 25 50 100 200 500 1000
L2-41
Free 25 50 100 200 500 1000
L2-42
Free 25 50 100 200 500 1000
L2-43
Free 25 50 100 200 500 1000

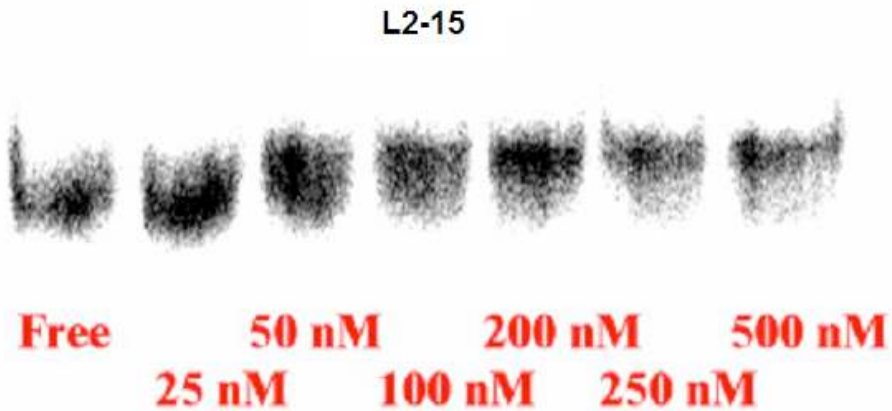
**Figure 83.** Gel band shift assay of mimetics from library **L2**. Peptides and **HIV TAR RNA** (~2 nM) were incubated at the peptide concentrations shown.

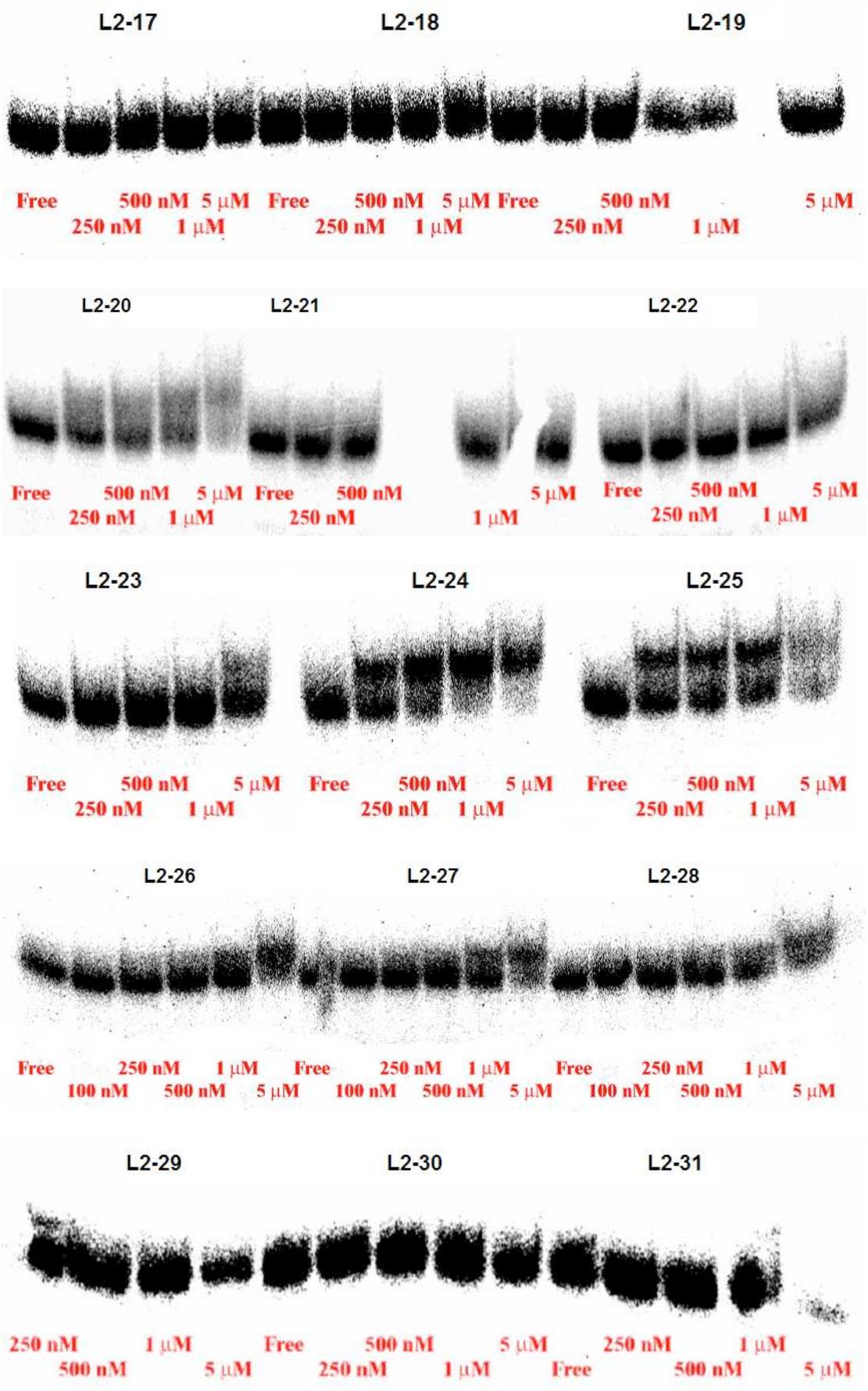


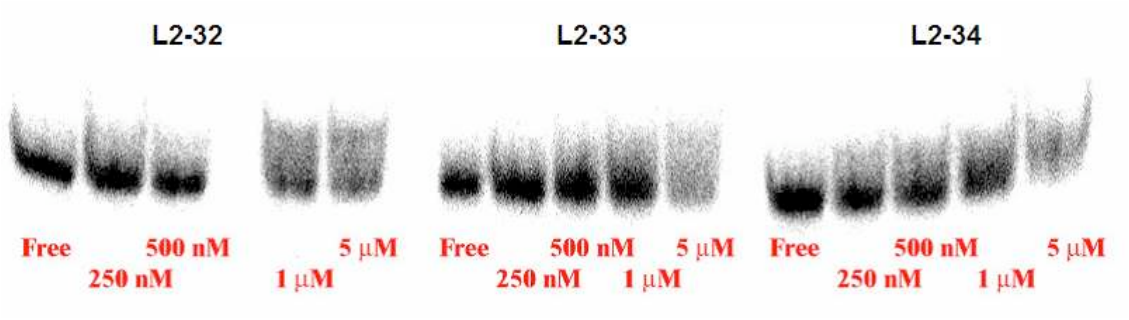

L2-08
Free
500 nM
1 μM
5 μM
L2-09
Free
500 nM
1 μM
5 μM
L2-10
Free
150 nM
500 nM
1 μM
5 μM

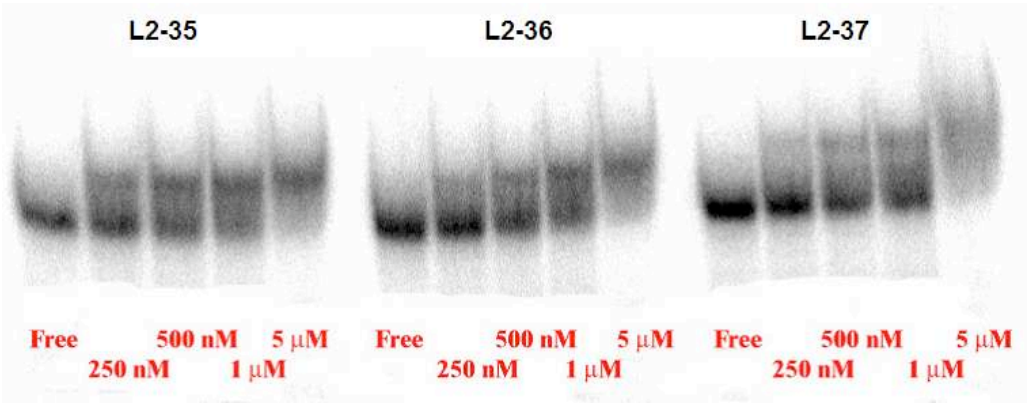

L2-11
Free
150 nM
500 nM
1 μM
5 μM
L2-12
Free
150 nM
500 nM
1 μM
5 μM
L2-13
Free
150 nM
500 nM
1 μM
5 μM

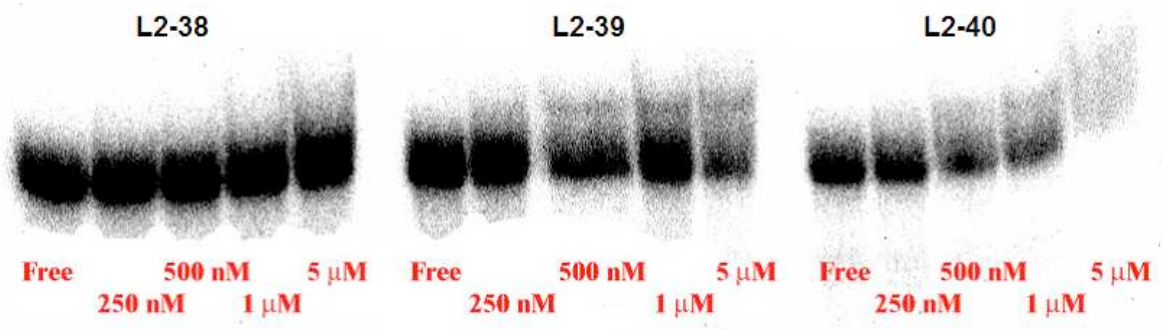

L2-14
Free
250 nM
500 nM
1 μM
5 μM
L2-15
Free
250 nM
500 nM
1 μM
5 μM
L2-16
Free
250 nM
500 nM
1 μM
5 μM


L2-15
Free
25 nM
50 nM
100 nM
200 nM
250 nM
500 nM


L2-17
L2-18
L2-19
Free
500 nM
5 μM
250 nM
1 μM
L2-20
L2-21
L2-22
L2-23
L2-24
L2-25
L2-26
L2-27
L2-28
100 nM
500 nM
L2-29
L2-30
L2-31


L2-32
Free
250 nM
500 nM
1 μM
5 μM
L2-33
Free
250 nM
500 nM
1 μM
5 μM
L2-34
Free
250 nM
500 nM
1 μM
5 μM


L2-35
Free
250 nM
500 nM
1 μM
5 μM
L2-36
Free
250 nM
500 nM
1 μM
5 μM
L2-37
Free
250 nM
500 nM
1 μM
5 μM


L2-38
Free
250 nM
500 nM
1 μM
5 μM
L2-39
Free
250 nM
500 nM
1 μM
5 μM
L2-40
Free
250 nM
500 nM
1 μM
5 μM

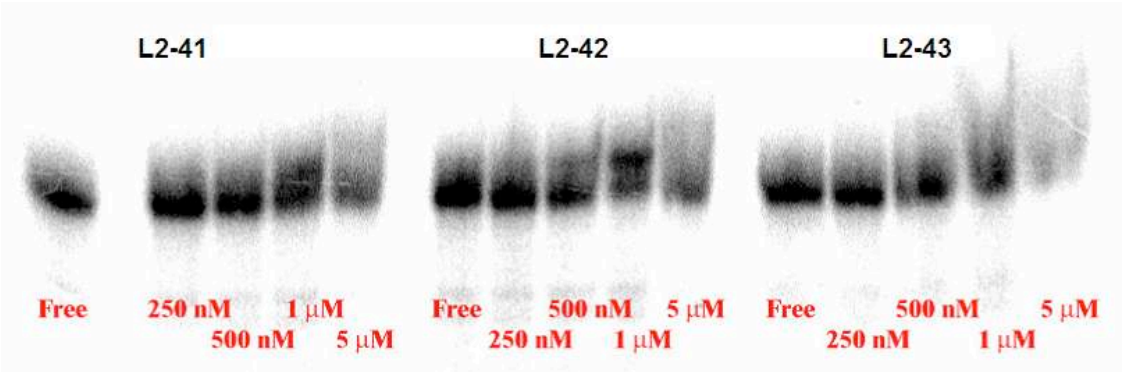

L2-41
Free
250 nM
500 nM
1 μM
5 μM
L2-42
Free
250 nM
500 nM
1 μM
5 μM
L2-43
Free
250 nM
500 nM
1 μM
5 μM

**Figure 84.** Inhibition assay. Inhibition of the complex formation between the ADP-1 peptide (150 nM) and HIV-1 TAR RNA (1 nM) by L2-15 at the peptide concentrations shown (nM and μM).

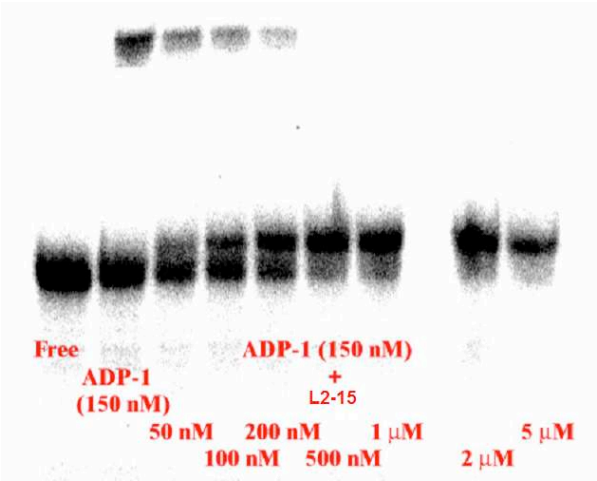



**Table 80.** Library **L3** of HIV Tat peptidomimetics. Residues 1 and 12 are attached to the D-Pro-L-Pro template.

| Mimetic | Position | | | | | | | | | | | | Template |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *1* | 2 | *3* | 4 | *5* | 6 | *7* | 8 | *9* | 10 | *11* | 12 | |
| **L3-01** | R | V | R | T | R | K | G | R | R | I | R | I | **D-Arg**-L-Pro |
| **L3-02** | R | V | R | T | R | K | G | R | R | **R** | R | I | D-Pro-L-Pro |
| **L3-03** | R | **C** | R | T | R | K | G | R | R | **R** | **C** | I | D-Pro-L-Pro |
| **L3-04** | R | V | R | T | R | **R** | G | R | R | I | R | I | D-Pro-L-Pro |
| **L3-05** | R | **C** | R | T | R | K | G | R | R | **K** | **C** | I | D-Pro-L-Pro |
| **L3-06** | R | V | R | T | R | **hR** | G | R | R | **K** | R | I | D-Pro-L-Pro |

**Figure 85.** Gel band shift assay of mimetics from library **L3**. Peptides and HIV TAR RNA (~2 nM) were incubated at the peptide concentrations shown.

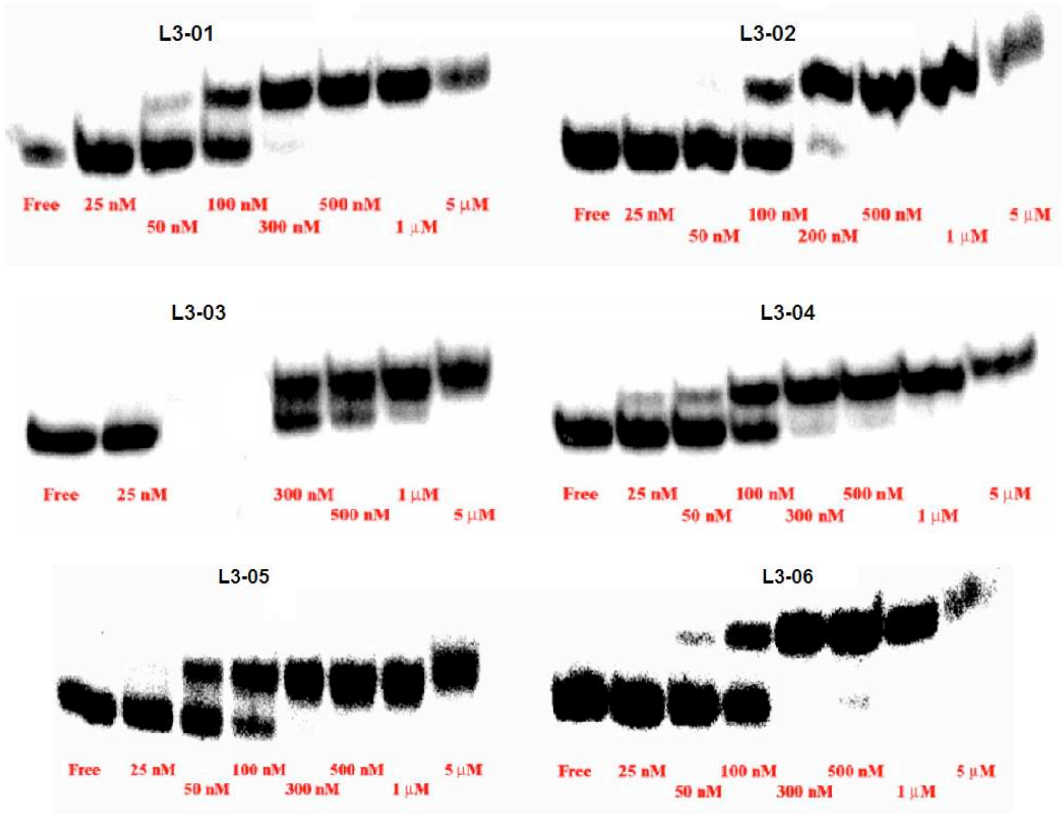
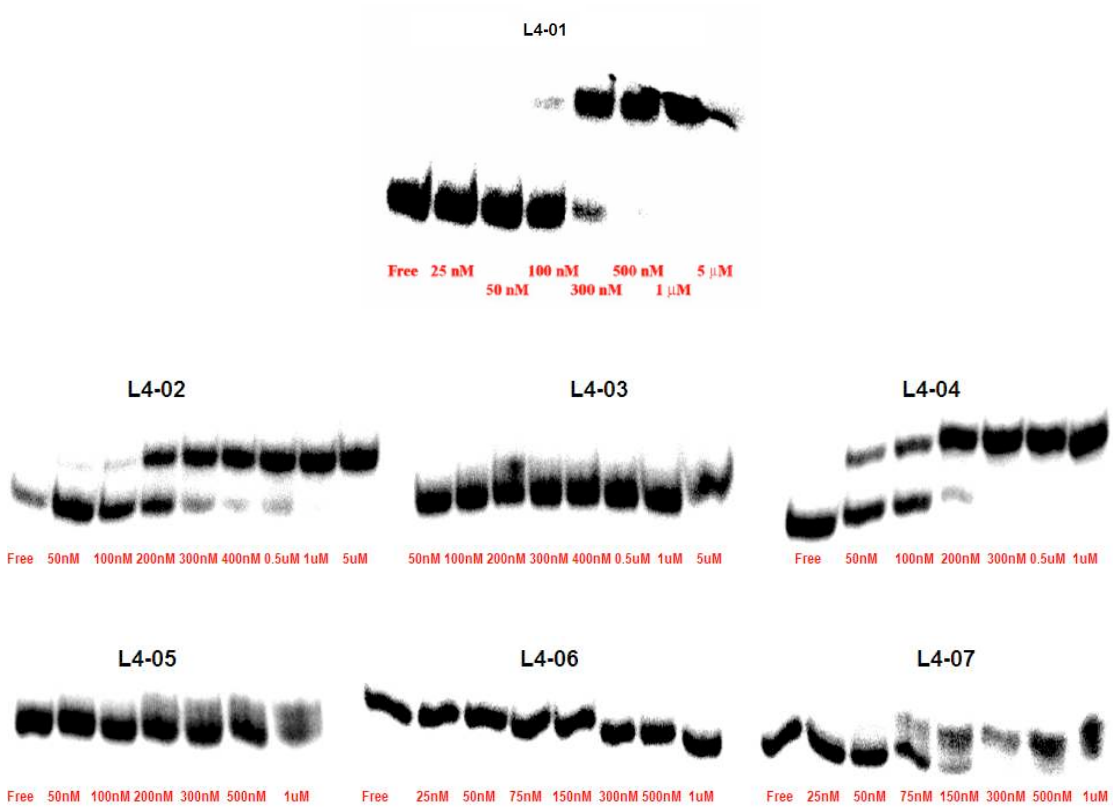


**Table 81.** Library **L4 (18-mer)** of HIV Tat peptidomimetics. Residues 1 and 16 are attached to the D-Pro-L-Pro template.

| **Mimetic** | **Position** | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | ***13*** | **14** | ***15*** | **16** |
| **L4-01** | **R** | **V** | **R** | ***C*** | **R** | **T** | **R** | **G** | **K** | **R** | **R** | **R** | ***C*** | **A** | **R** | **V** |
| **L4-02** | R | V | R | *C* | R | Q | R | K | G | R | A | R | *C* | I | R | I |
| **L4-03** | R | V | R | *C* | R | K | L | K | G | Q | T | R | *C* | I | R | I |
| **L4-04** | R | V | R | *C* | R | K | R | $^{D}$R | G | Q | T | R | *C* | I | R | I |
| **L4-05** | R | V | R | *C* | R | Q | R | G | P | G | K | R | *C* | I | R | I |
| **L4-06** | R | V | R | *C* | R | K | Q | G | P | G | T | R | *C* | I | R | I |
| **L4-07** | R | V | R | *C* | R | R | K | G | P | G | Q | R | *C* | I | R | I |
| **L4-08** | R | V | R | *C* | R | K | R | G | P | G | Q | R | *C* | I | R | I |
| **L4-09** | R | *C* | R | V | R | Q | R | K | G | R | A | R | K | I | *C* | I |
| **L4-10** | R | *C* | R | V | R | K | Q | G | P | G | T | R | K | I | *C* | I |
| **L4-11** | R | *C* | R | V | R | K | Q | G | P | G | K | R | K | I | *C* | I |
| **L4-12** | R | *C* | R | V | R | K | R | G | P | G | Q | R | K | I | *C* | I |
| **L4-13** | R | *C* | R | V | R | R | K | G | P | G | A | R | K | I | *C* | I |

**Figure 86.** Gel band shift assay of mimetics from library **L4 (18-mer)**. Peptides and HIV TAR RNA (~2 nM) were incubated at the peptide concentrations shown.



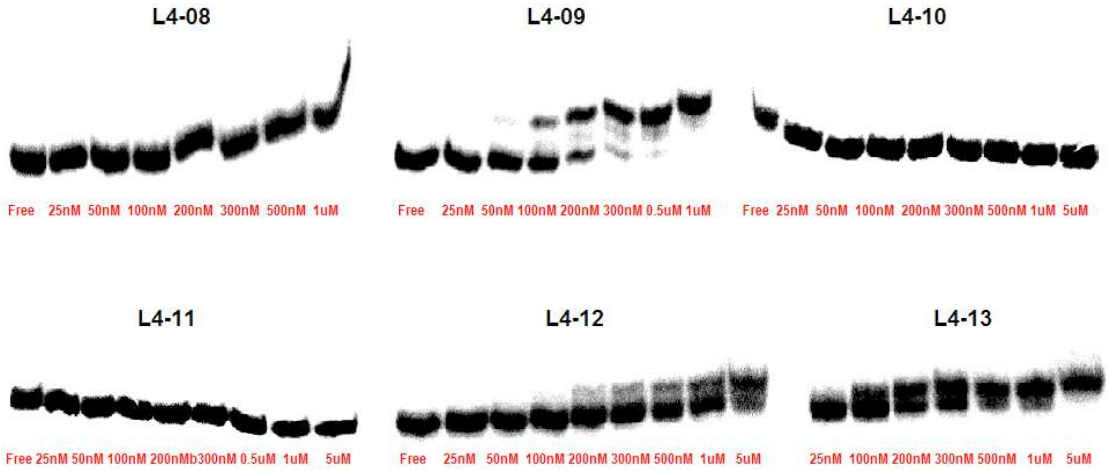



**Table 82.** Library **L4 (15-mer)** of HIV Tat peptidomimetics. Residues 1 and 13 are attached to the D-Pro-L-Pro template.

| **Mimetic** | **Position** | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | ***13*** |
| **L4-14** | R | V | R | C | R | K | K | G | K | C | I | R | I |
| **L4-15** | R | V | R | C | R | R | R | G | R | C | I | R | I |
| **L4-16** | R | V | R | T | R | K | K | G | K | R | I | R | I |
| **L4-17** | R | V | R | C | R | K | K | $^{D}$P | K | C | I | R | I |
| **L4-18** | R | V | R | T | R | K | K | $^{D}$P | K | R | I | R | I |
| **L4-19** | R | V | R | V | R | K | K | G | R | T | I | R | I |

**Figure 87.** Gel band shift assay of mimetics from library **L4 (15-mer)**. Peptides and HIV TAR RNA (~2 nM) were incubated at the peptide concentrations shown.

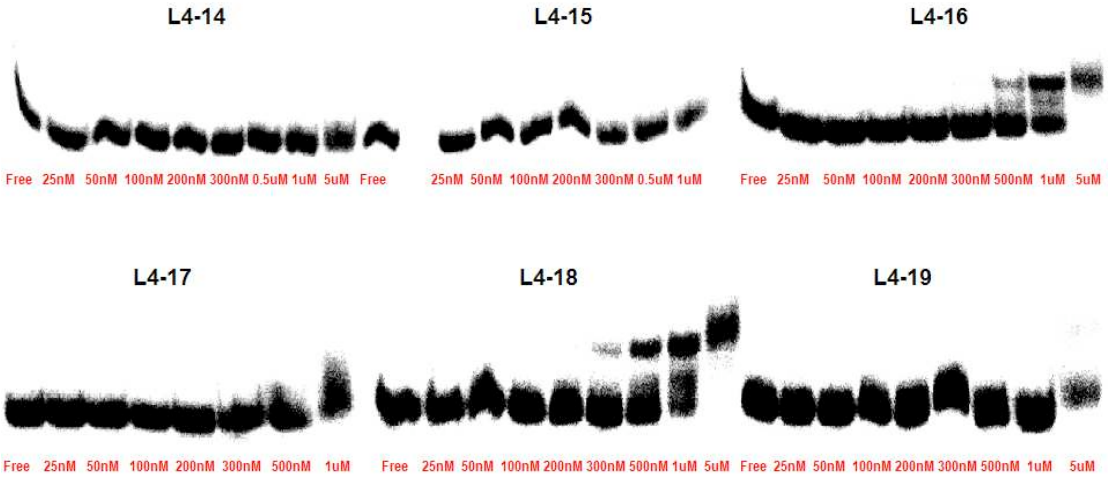


**Table 83.** Library **L4 (16-mer)** of HIV Tat peptidomimetics. Residues 1 and 14 are attached to the D-Pro-L-Pro template.

| **Mimetic** | **Position** | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | ***13*** | **14** |
| **L4-20** | R | V | R | C | R | K | R | K | G | R | C | I | R | I |
| **L4-21** | R | V | R | C | R | R | K | R | G | R | C | I | R | I |
| **L4-22** | R | V | R | T | R | Q | R | K | G | R | R | I | R | I |
| **L4-23** | R | V | R | V | R | K | R | K | G | R | Q | I | R | I |
| **L4-24** | R | V | R | Q | R | K | R | K | G | R | V | I | R | I |
| **L4-25** | R | V | R | C | R | K | R | K | $^{D}$P | R | C | I | R | I |

**Figure 88.** Gel band shift assay of mimetics from library **L4 (16-mer)**. Peptides and HIV TAR RNA (~2 nM) were incubated at the peptide concentrations shown.

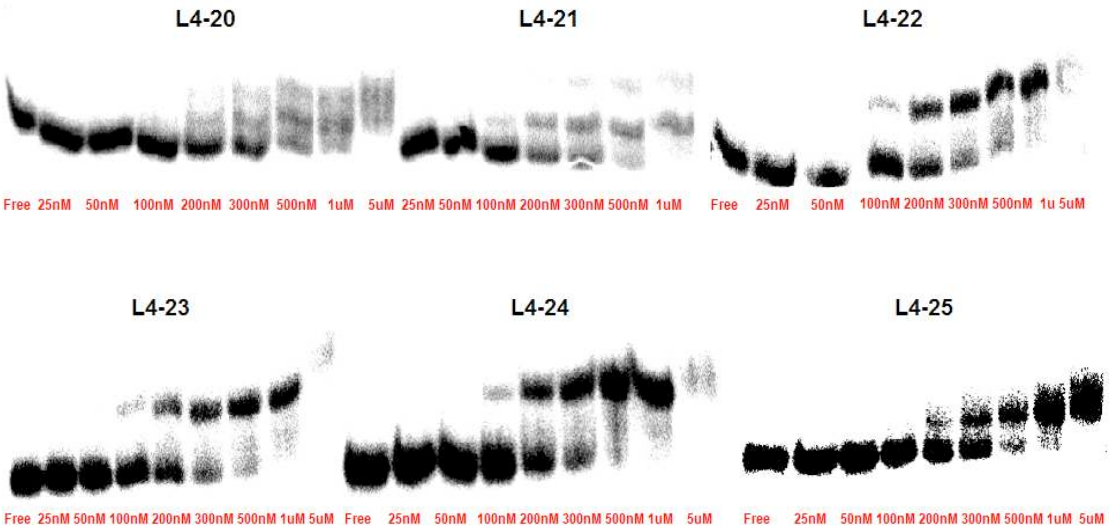


**Table 84.** Library **L4-04K** of HIV Tat mimetics. Residues 1 and 16 are attached to the D-Pro-L-Pro template.

| Mimetic | Position | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *1* | 2 | *3* | 4 | *5* | 6 | *7* | 8 | *9* | 10 | *11* | 12 | *13* | 14 | *15* | 16 |
| **L4-04** | **R** | **V** | **R** | **C** | **R** | **K** | **R** | **$^{D}$R** | **G** | **Q** | **T** | **R** | **C** | **I** | **R** | **I** |
| **L4-04K01** | **K** | V | R | C | R | K | R | $^{D}$R | G | Q | T | R | C | I | R | I |
| **L4-04K02** | R | V | **K** | C | R | K | R | $^{D}$R | G | Q | T | R | C | I | R | I |
| **L4-04K03** | R | V | R | C | **K** | K | R | $^{D}$R | G | Q | T | R | C | I | R | I |
| **L4-04K04** | R | V | R | C | R | K | **K** | $^{D}$R | G | Q | T | R | C | I | R | I |
| **L4-04K05** | R | V | R | C | R | K | R | $^{D}$R | G | Q | T | **K** | C | I | R | I |
| **L4-04K06** | R | V | R | C | R | K | R | $^{D}$R | G | Q | T | R | C | I | **K** | I |
| **L4-04K07** | R | V | R | C | R | K | R | **$^{D}$K** | G | Q | T | R | C | I | R | I |

**Figure 89.** Gel band shift assay of mimetics from library **L4-04K**. Peptides and HIV TAR RNA (~2 nM) were incubated at the peptide concentrations shown.

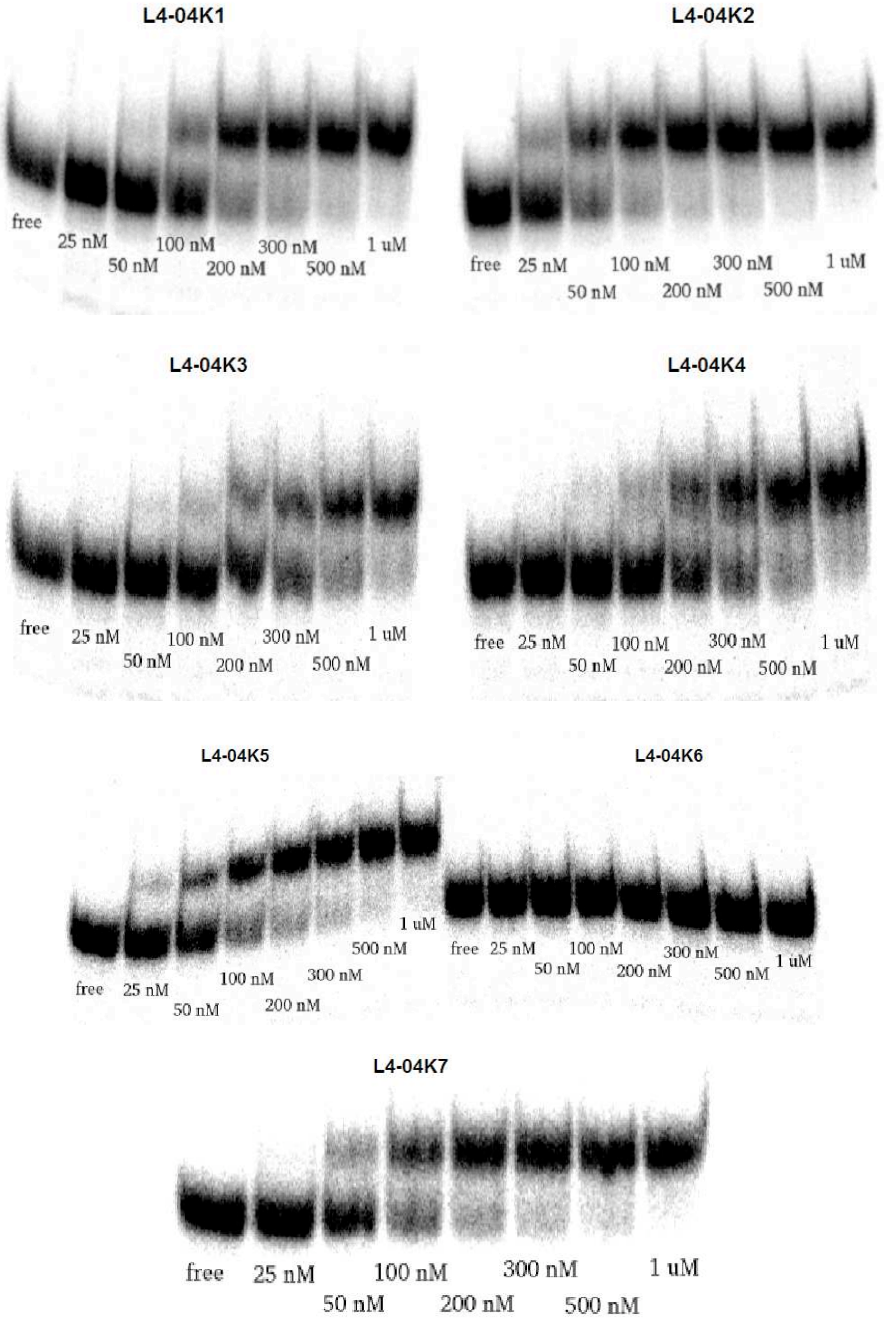


**Table 85.** Library **L2-15hR** of HIV Tat mimetics. Residues 1 and 12 are attached to the D-Pro-L-Pro template.

| **Mimetic** | **Position** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** |
| **L2-15** | **R** | **V** | **R** | **T** | **R** | **K** | **G** | **R** | **R** | **I** | **R** | **I** |
| **L2-15hR01** | **hR** | V | R | T | R | K | G | R | R | I | R | I |
| **L2-15hR02** | R | V | **hR** | T | R | K | G | R | R | I | R | I |
| **L2-15hR03** | R | V | R | T | **hR** | K | G | R | R | I | R | I |
| **L2-15hR04** | R | V | R | T | R | K | G | **hR** | R | I | R | I |
| **L2-15hR05** | R | V | R | T | R | K | G | R | **hR** | I | R | I |
| **L2-15hR06** | R | V | R | T | R | K | G | R | R | I | **hR** | I |

**Figure 90.** Gel band shift assay of mimetics from library **L2-15hR**. Peptides and HIV TAR RNA (~2 nM) were incubated at the peptide concentrations shown.

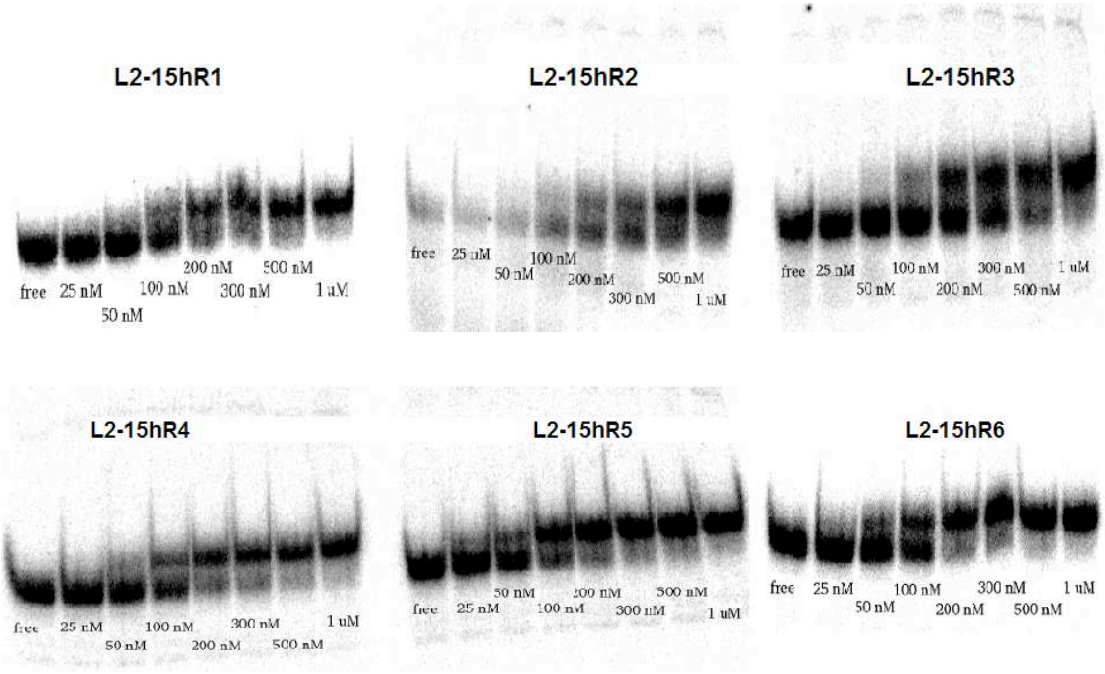


**Table 86.** Derivatives of L4-04 of HIV Tat mimetics.

| **Mimetic** | **Position** | | | | | | | | | | | | | | | **Template** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | ***13*** | **14** | ***15*** | **16** | |
| **L4-04** | **R** | **V** | **R** | ***C*** | **R** | **K** | **R** | **$^{D}$R** | **G** | **Q** | **T** | **R** | ***C*** | **I** | **R** | **I** | D-Pro-L-Pro |
| **L4-04a** | **Ac-R** | V | R | *C* | R | K | R | $^{D}$R | G | Q | T | R | *C* | I | R | **I-$NH_2$** | -- |
| **L4-04b** | R | V | R | **Abu** | R | K | R | $^{D}$R | G | Q | T | R | **Abu** | I | R | I | D-Pro-L-Pro |
| **L4-04c** | **Ac-R** | V | R | **Abu** | R | K | R | $^{D}$R | G | Q | T | R | **Abu** | I | R | **I-$NH_2$** | -- |

**Figure 91.** Gel band shift assay of derivates of L4-04 peptide. Peptides and HIV TAR RNA (~2 nM) were incubated at the peptide concentrations shown.

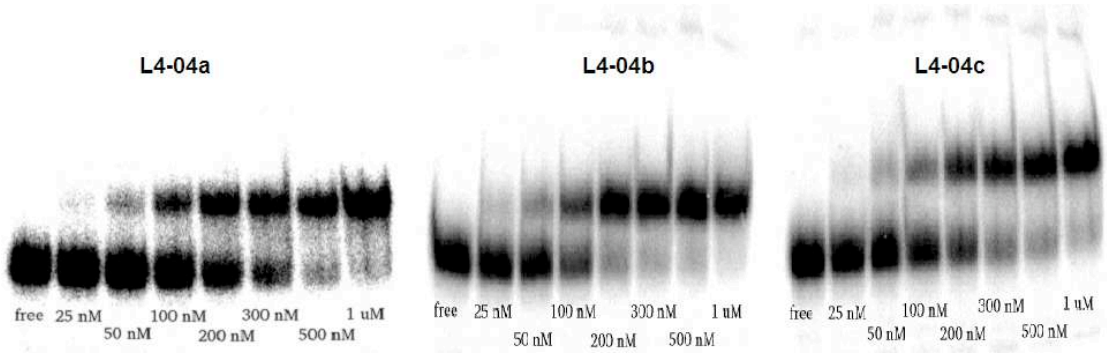


**Table 87.** Library **L5** of BIV Tat mimetics. Residues 1 and 16 are attached to the D-Pro-L-Pro template.

| Mimetic | Position | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | ***13*** | **14** | ***15*** | **16** |
| **L5-01** | R | V | R | *C* | R | K | R | R | G | Q | T | R | *C* | I | R | I |
| **L5-02** | R | V | K | *C* | R | K | R | K | G | Q | T | R | *C* | I | R | I |
| **L5-03** | R | V | R | *C* | R | K | L | $^{D}$R | G | Q | T | R | *C* | I | R | I |
| **L5-04** | R | V | R | *C* | R | R | K | $^{D}$R | G | Q | T | R | *C* | I | R | I |
| **L5-05** | R | V | R | T | R | K | R | $^{D}$R | G | Q | T | R | A | I | K | I |
| **L5-06** | R | *C* | R | V | R | K | R | $^{D}$R | G | Q | T | R | R | I | *C* | I |
| **L5-07** | R | V | R | *C* | R | K | R | G | P | G | A | R | *C* | I | K | I |
| **L5-08** | R | *C* | R | V | R | R | K | K | G | R | A | R | K | I | *C* | I |
| **L5-09** | R | *C* | R | V | R | Q | R | G | P | G | A | R | K | I | *C* | I |
| **L5-10** | R | V | R | *C* | R | T | R | K | G | R | A | R | *C* | I | K | I |
| **L5-11** | R | V | R | *C* | R | Q | R | K | G | R | R | A | *C* | I | R | I |
| **L5-12** | R | V | R | *C* | R | Q | R | K | G | R | R | I | *C* | I | R | I |

**Figure 92.** Gel band shift assay of mimetics from library **L5**. Peptides and HIV TAR RNA (~2 nM) were incubated at the peptide concentrations shown.

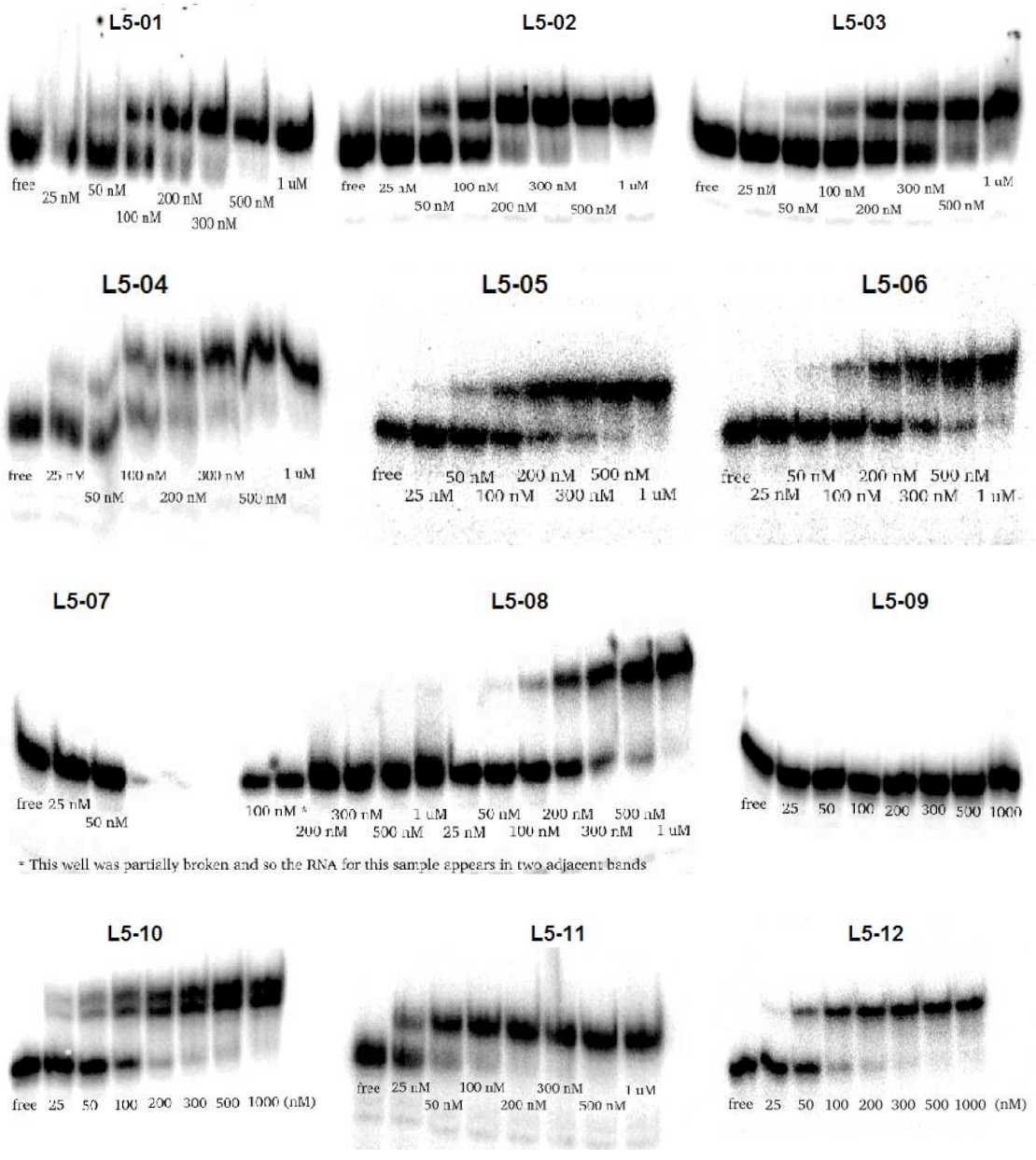


**Figure 93.** Gel band shift assay of selected mimetics. Peptides and HIV TAR RNA (~2 nM) were incubated at the peptide concentrations shown. **EMSA was performed in the absence of tRNA.**

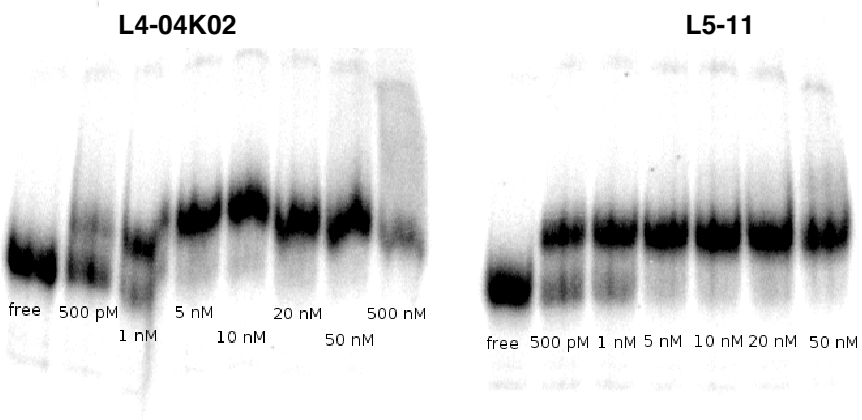



**Table 88.** Library **L4-04hR** of HIV Tat mimetics. Residues 1 and 16 are attached to the D-Pro-L-Pro template.

| **Mimetic** | **Position** | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** | ***13*** | **14** | ***15*** | **16** |
| **L4-04** | **R** | **V** | **R** | ***C*** | **R** | **K** | **R** | **$^{D}$R** | **G** | **Q** | **T** | **R** | ***C*** | **I** | **R** | **I** |
| **L4-04hR01** | **hR** | V | R | *C* | R | K | R | $^{D}$R | G | Q | T | R | *C* | I | R | I |
| **L4-04hR02** | R | V | **hR** | *C* | R | K | R | $^{D}$R | G | Q | T | R | *C* | I | R | I |
| **L4-04hR03** | R | V | R | *C* | **hR** | K | R | $^{D}$R | G | Q | T | R | *C* | I | R | I |
| **L4-04hR04** | R | V | R | *C* | R | K | **hR** | $^{D}$R | G | Q | T | R | *C* | I | R | I |
| **L4-04hR05** | R | V | R | *C* | R | K | R | $^{D}$R | G | Q | T | **hR** | *C* | I | R | I |
| **L4-04hR06** | **R** | V | R | *C* | R | K | R | $^{D}$R | G | Q | T | R | *C* | I | **hR** | I |

**Figure 94.** Gel band shift assay of mimetics from library **L4-04hR**. Peptides and HIV TAR RNA (~2 nM) were incubated at the peptide concentrations shown.

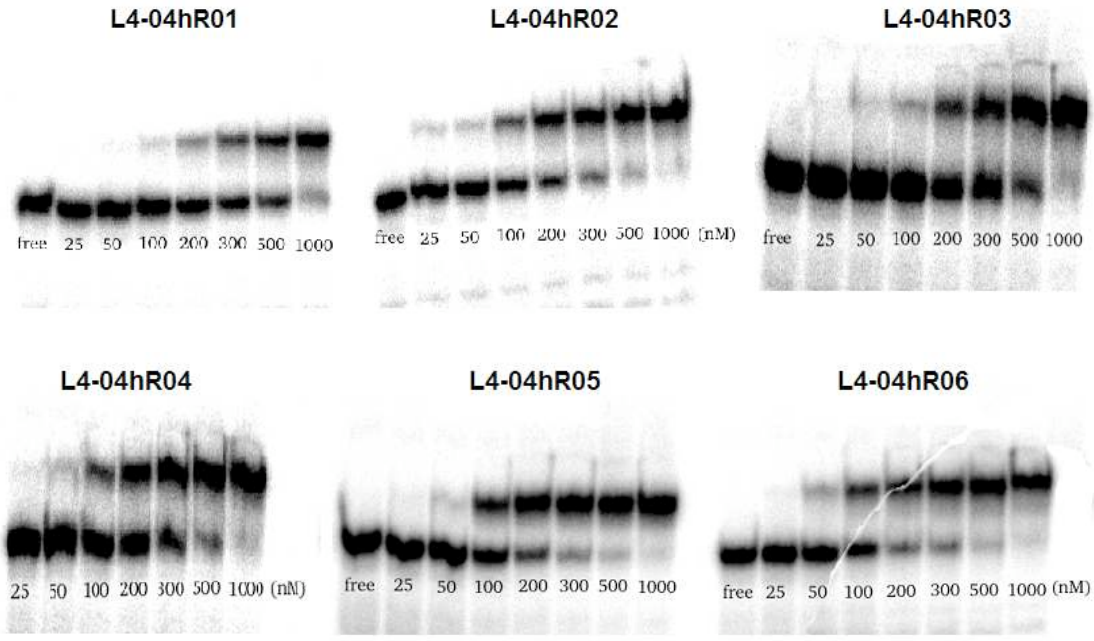
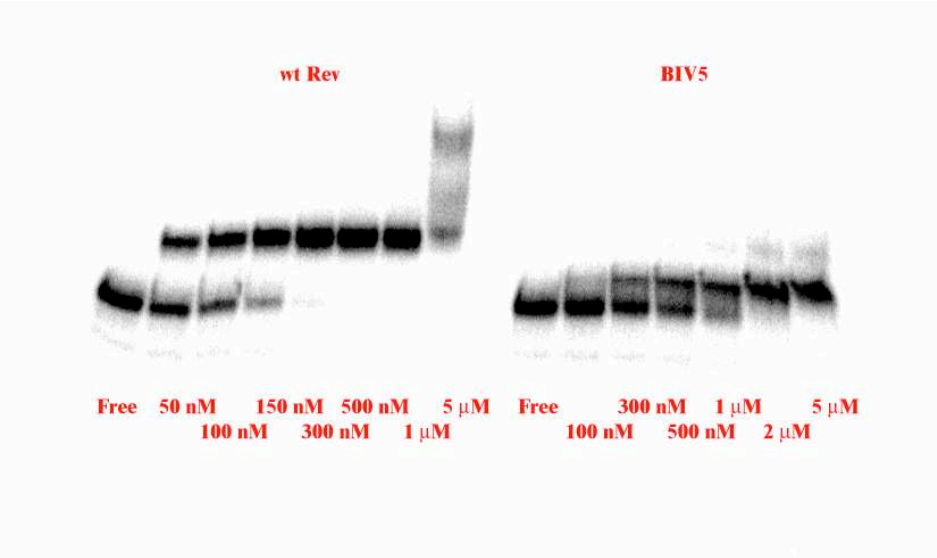
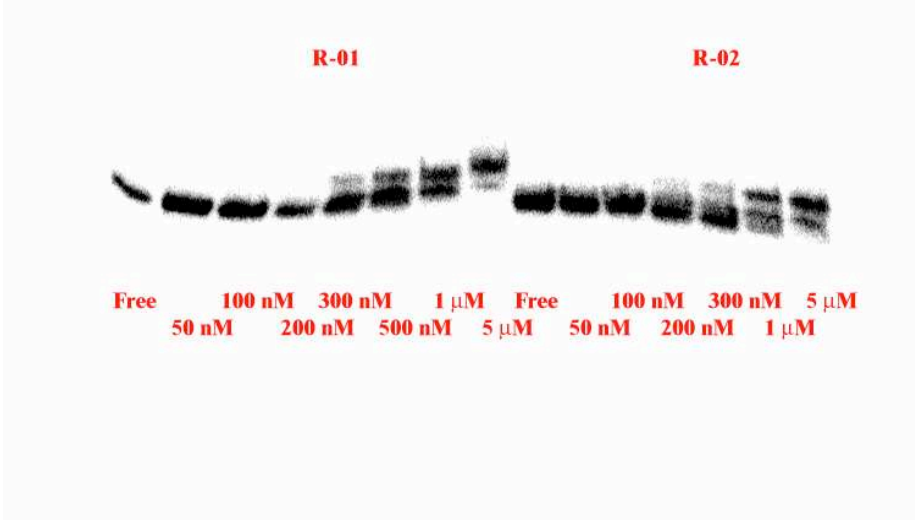


**Table 89.** Library **R1** of Rev peptidomimetics. Residues 1 and 12 are attached to the D-Pro-L-Pro template.

| **Mimetic** | **Position** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** |
| **Rev** | **Ac-RQARRNRRRRWRERNR-$NH_2$** | | | | | | | | | | | |
| **BIV5** | R | R | G | T | R | G | K | R | R | I | G | R |
| **R1-01** | W | R | R | R | A | T | $^{D}$R | Q | R | N | R | R |
| **R1-02** | W | R | R | R | A | P | $^{D}$R | Q | R | N | R | R |
| **R1-03** | W | R | R | R | G | P | $^{D}$R | Q | R | N | R | R |
| **R1-04** | W | R | R | R | V | P | $^{D}$R | Q | R | N | R | R |
| **R1-05** | W | R | R | R | A | G | $^{D}$R | Q | R | N | R | R |
| **R1-06** | W | R | R | R | A | G | K | Q | R | N | R | R |
| **R1-07** | W | R | R | R | A | T | $^{D}$R | G | R | N | R | R |
| **R1-08** | W | R | R | R | A | P | $^{D}$R | G | R | N | R | R |
| **R1-09** | W | R | R | R | G | P | $^{D}$R | G | R | N | R | R |
| **R1-10** | W | R | R | R | G | T | $^{D}$R | Q | R | N | R | R |
| **R1-11** | W | R | R | R | V | G | $^{D}$R | Q | R | N | R | R |
| **R1-12** | W | R | R | R | A | S | $^{D}$R | Q | R | N | R | R |

**Figure 95.** Gel band shift assay of mimetics from library **R1**. Peptides and HIV RRE RNA (~2 nM) were incubated at the peptide concentrations shown.



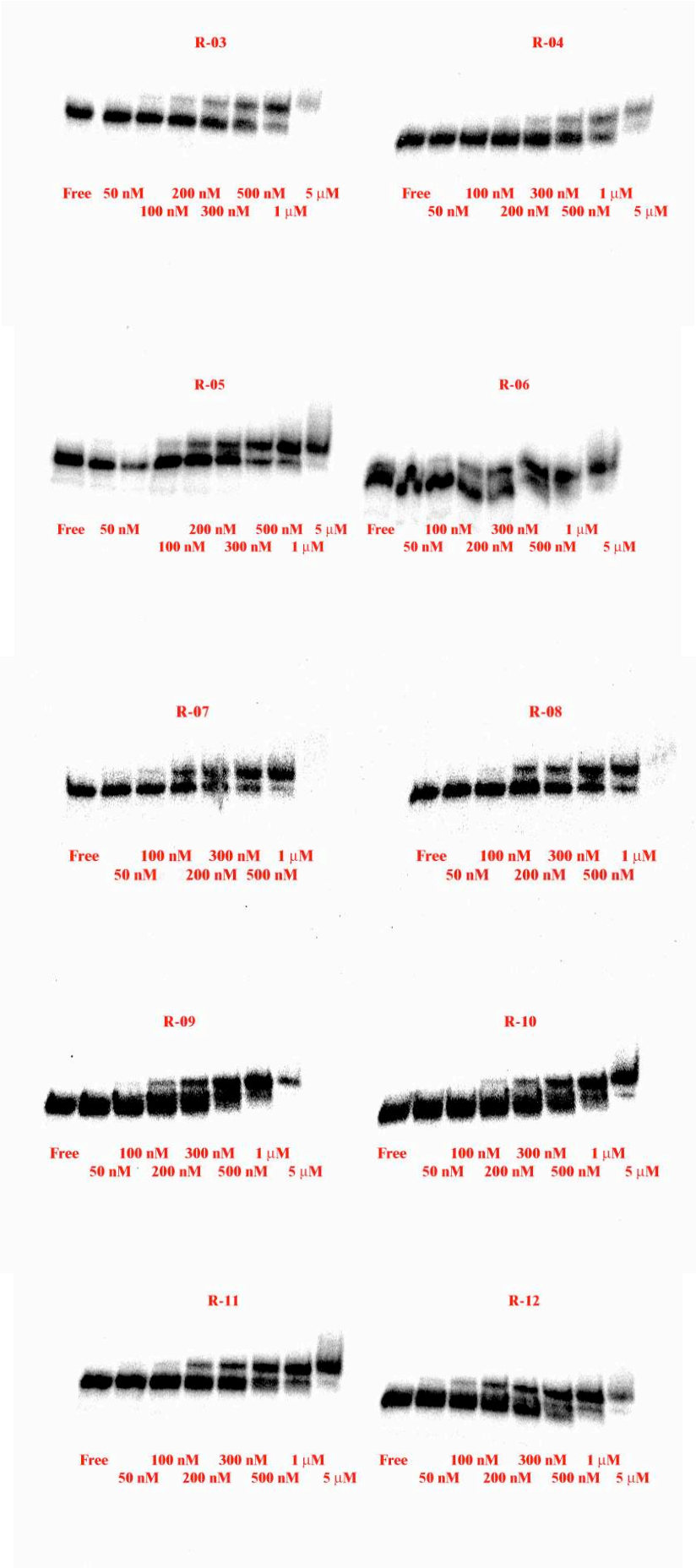
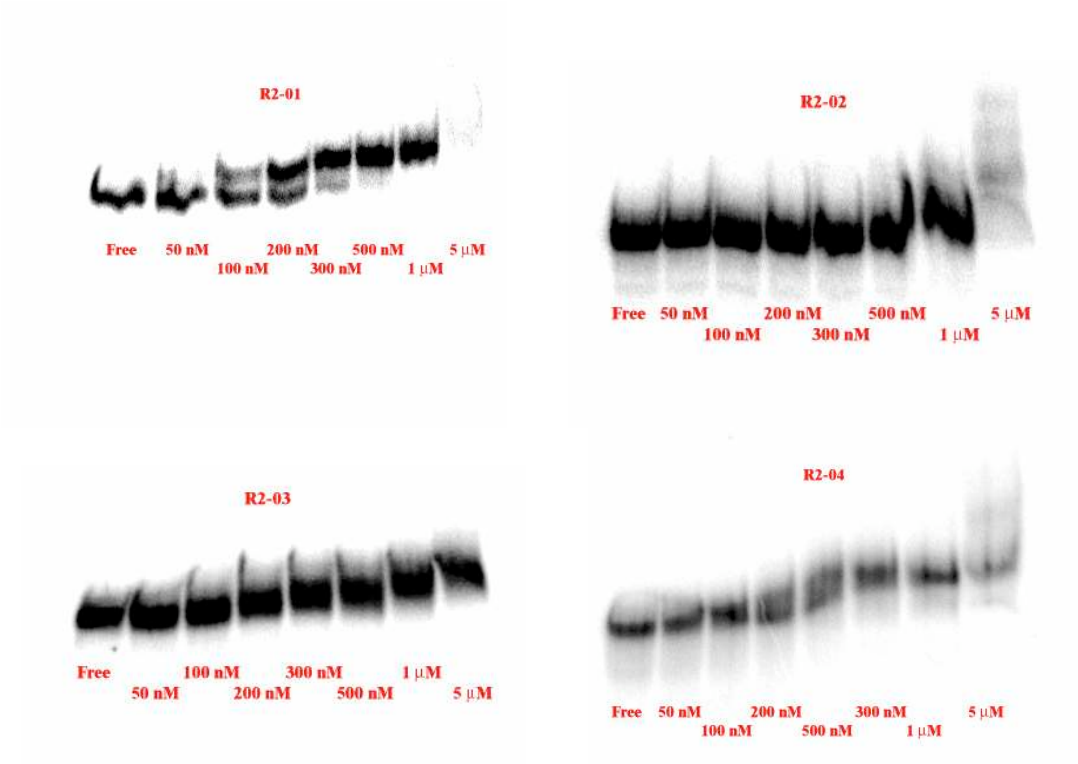
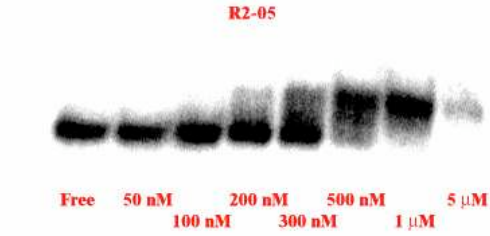

R-03
Free 50 nM 200 nM 500 nM 5 μM
100 nM 300 nM 1 μM
R-04
Free 100 nM 300 nM 1 μM
50 nM 200 nM 500 nM 5 μM
R-05
Free 50 nM 200 nM 500 nM 5 μM
100 nM 300 nM 1 μM
R-06
Free 100 nM 300 nM 1 μM
50 nM 200 nM 500 nM 5 μM
R-07
Free 100 nM 300 nM 1 μM
50 nM 200 nM 500 nM
R-08
Free 100 nM 300 nM 1 μM
50 nM 200 nM 500 nM
R-09
Free 100 nM 300 nM 1 μM
50 nM 200 nM 500 nM 5 μM
R-10
Free 100 nM 300 nM 1 μM
50 nM 200 nM 500 nM 5 μM
R-11
Free 100 nM 300 nM 1 μM
50 nM 200 nM 500 nM 5 μM
R-12
Free 100 nM 300 nM 1 μM
50 nM 200 nM 500 nM 5 μM

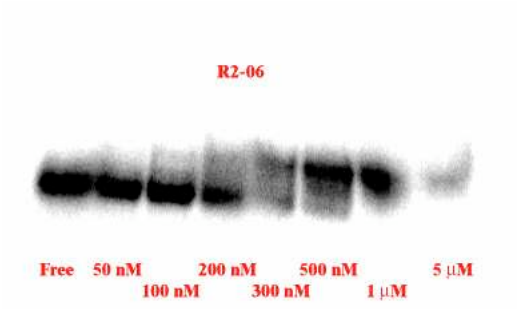
**Table 90.** Library **R2** of Rev peptidomimetics. Residues 1 and 12 are attached to the D-Pro-L-Pro template.

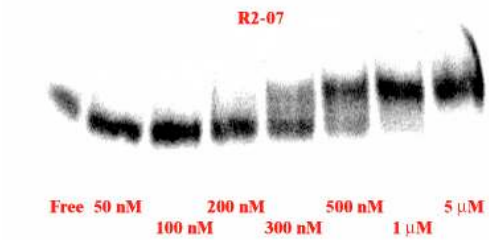
| **Mimetic** | **Position** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***1*** | **2** | ***3*** | **4** | ***5*** | **6** | ***7*** | **8** | ***9*** | **10** | ***11*** | **12** |
| **R2-01** | W | R | R | R | A | K | G | Q | R | N | R | R |
| **R2-02** | W | R | L | R | A | K | G | Q | R | N | R | R |
| **R2-03** | W | R | Q | R | A | K | G | Q | R | N | R | R |
| **R2-04** | W | R | R | R | A | K | G | Q | R | N | R | V |
| **R2-05** | W | R | L | R | A | $^{D}$R | K | G | R | N | R | R |
| **R2-06** | W | R | L | R | A | K | K | G | R | N | R | R |
| **R2-07** | W | R | K | R | A | G | K | Q | R | N | R | R |
| **R2-08** | W | R | K | R | A | K | G | Q | R | N | R | R |
| **R2-09** | W | R | R | R | A | $^{D}$R | G | Q | R | N | R | R |
| **R2-10** | W | R | L | R | A | $^{D}$R | G | K | R | N | R | R |
| **R2-11** | R | R | L | R | A | $^{D}$R | G | Q | R | N | R | R |
| **R2-12** | R | R | L | R | A | K | G | Q | R | N | R | R |
| **R2-13** | R | V | R | R | A | K | G | Q | R | N | R | R |
| **R2-14** | R | C | R | R | A | K | G | Q | R | R | C | R |
| **R2-15** | K | R | Q | R | T | K | G | R | R | L | O | R |
| **R2-16** | W | R | R | R | A | G | R | Q | R | N | R | R |

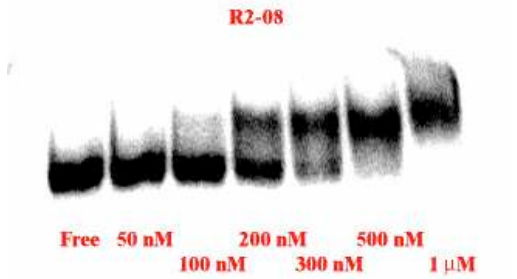
**Figure 96.** Gel band shift assay of mimetics from library **R2**. Peptides and HIV RRE RNA (~2 nM) were incubated at the peptide concentrations shown.

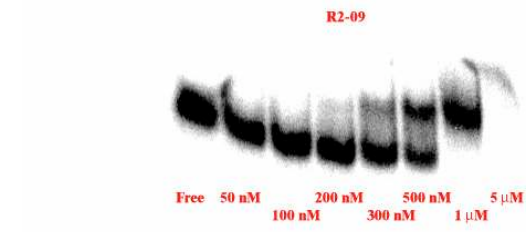


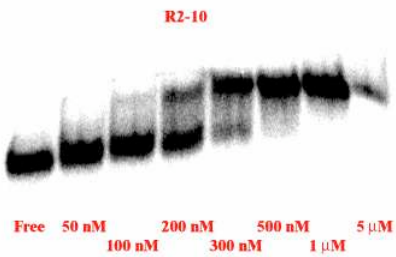

R2-05
Free
50 nM
100 nM
200 nM
300 nM
500 nM
1 µM
5 µM

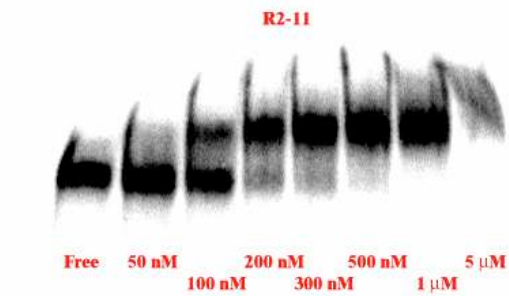

R2-06
Free
50 nM
100 nM
200 nM
300 nM
500 nM
1 µM
5 µM

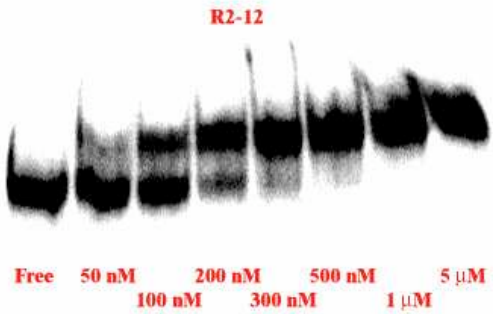

R2-07
Free
50 nM
100 nM
200 nM
300 nM
500 nM
1 µM
5 µM


R2-08
Free
50 nM
100 nM
200 nM
300 nM
500 nM
1 µM


R2-09
Free
50 nM
100 nM
200 nM
300 nM
500 nM
1 µM
5 µM


R2-10
Free
50 nM
100 nM
200 nM
300 nM
500 nM
1 µM
5 µM


R2-11
Free
50 nM
100 nM
200 nM
300 nM
500 nM
1 µM
5 µM


R2-12
Free
50 nM
100 nM
200 nM
300 nM
500 nM
1 µM
5 µM

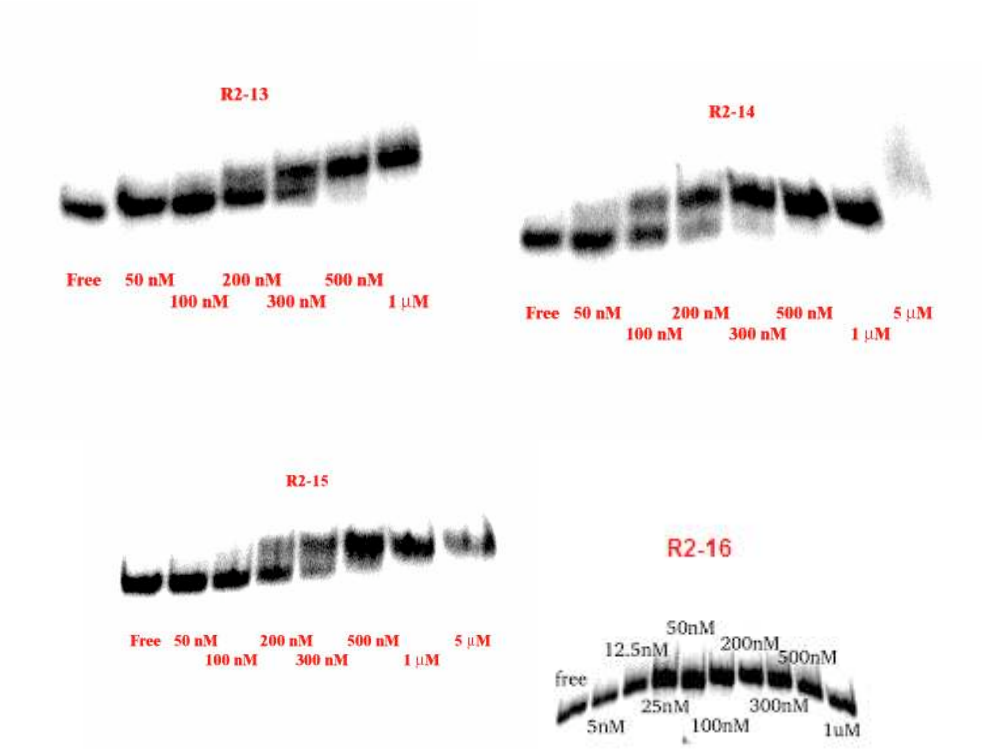
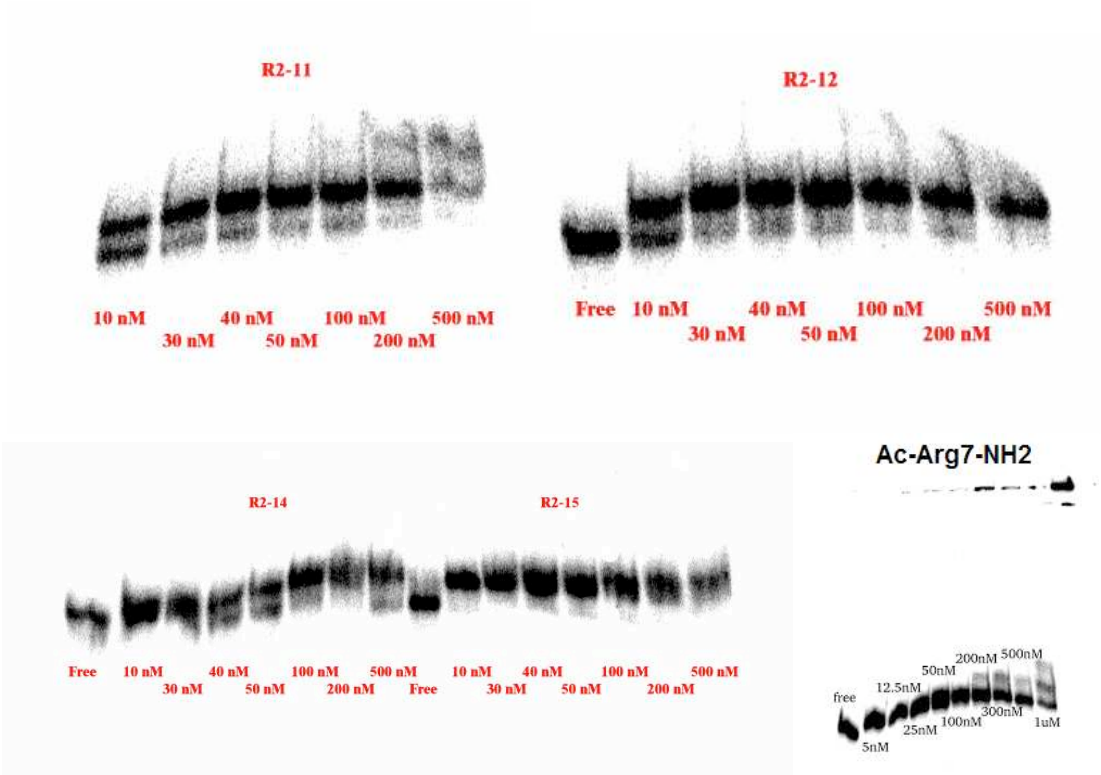



**Figure 97.** Gel band shift assay of selected Rev peptidomimetics and Ac-$Arg_7$-$NH_2$. Peptides and HIV RRE RNA (~2 nM) were incubated at the peptide concentrations shown. **EMSA was performed in the absence of tRNA.**



# 7. BIBLIOGRAPHY


(1) Silva, J. G.; Carvalho, I. New insights into aminoglycoside antibiotics and derivatives. *Curr Med Chem* **2007**, *14*, 1101-1119.

(2) Dykxhoorn, D. M.; Lieberman, J. Running interference: prospects and obstacles to using small interfering RNAs as small molecule drugs. *Annu Rev Biomed Eng* **2006**, *8*, 377-402.

(3) Platt, T. Rho and RNA: models for recognition and response. *Mol Microbiol* **1994**, *11*, 983-990.

(4) Henkin, T. M.; Yanofsky, C. Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions. *Bioessays* **2002**, *24*, 700-707.

(5) Brady, J.; Kashanchi, F. Tat gets the "green" light on transcription initiation. *Retrovirology* **2005**, *2*, 69.

(6) Day, D. A.; Tuite, M. F. Post-transcriptional gene regulatory mechanisms in eukaryotes: an overview. *J Endocrinol* **1998**, *157*, 361-371.

(7) Krecic, A. M.; Swanson, M. S. hnRNP complexes: composition, structure, and function. *Curr Opin Cell Biol* **1999**, *11*, 363-371.

(8) Butcher, S. E.; Brow, D. A. Towards understanding the catalytic core structure of the spliceosome. *Biochem Soc Trans* **2005**, *33*, 447-449.

(9) Burd, C. G.; Dreyfuss, G. Conserved structures and diversity of functions of RNA-binding proteins. *Science* **1994**, *265*, 615-621.

(10) Siomi, H.; Dreyfuss, G. RNA-binding proteins as regulators of gene expression. *Curr Opin Genet Dev* **1997**, *7*, 345-353.

(11) Maris, C.; Dominguez, C.; Allain, F. H. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *Febs J* **2005**, *272*, 2118-2131.

(12) Oubridge, C.; Ito, N.; Evans, P. R.; Teo, C. H.; Nagai, K. Crystal structure at 1.92 A resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **1994**, *372*, 432-438.

(13) Musco, G.; Stier, G.; Joseph, C.; Castiglione Morelli, M. A.; Nilges, M. et al. Three-dimensional structure and stability of the KH domain: molecular insights into the fragile X syndrome. *Cell* **1996**, *85*, 237-245.

(14) Grishin, N. V. KH domain: one motif, two folds. *Nucl Acids Res* **2001**, *29*, 638-643.

(15) Doyle, M.; Jantsch, M. F. New and old roles of the double-stranded RNA-binding domain. *J Struct Biol* **2002**, *140*, 147-153.

(16) Gamsjaeger, R.; Liew, C. K.; Loughlin, F. E.; Crossley, M.; Mackay, J. P. Sticky fingers: zinc-fingers as protein-recognition motifs. *Trends Biochem Sci* **2007**, *32*, 63-70.

(17) Brown, R. S. Zinc finger proteins: getting a grip on RNA. *Curr Opin Struct Biol* **2005**, *15*, 94-98.

(18) Lewis, H. A.; Musunuru, K.; Jensen, K. B.; Edo, C.; Chen, H. et al. Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. *Cell* **2000**, *100*, 323-332.

(19) Wu, H.; Henras, A.; Chanfreau, G.; Feigon, J. Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. *Proc Natl Acad Sci U S A* **2004**, *101*, 8307-8312.

(20) Hudson, B. P.; Martinez-Yamout, M. A.; Dyson, H. J.; Wright, P. E. Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. *Nat Struct Mol Biol* **2004**, *11*, 257-264.
(21) Scharpf, M.; Sticht, H.; Schweimer, K.; Boehm, M.; Hoffmann, S. et al. Antitermination in bacteriophage lambda. The structure of the N36 peptide-boxB RNA complex. *Eur J Biochem* **2000**, *267*, 2397-2408.
(22) Xia, T.; Wan, C.; Roberts, R. W.; Zewail, A. H. RNA-protein recognition: single-residue ultrafast dynamical control of structural specificity and function. *Proc Natl Acad Sci U S A* **2005**, *102*, 13013-13018.
(23) Hemmerich, P.; Bosbach, S.; von Mikecz, A.; Krawinkel, U. Human ribosomal protein L7 binds RNA with an alpha-helical arginine-rich and lysine-rich domain. *Eur J Biochem* **1997**, *245*, 549-556.
(24) Hancock, R. E. Cationic peptides: effectors in innate immunity and novel antimicrobials. *Lancet Infect Dis* **2001**, *1*, 156-164.
(25) Battiste, J. L.; Mao, H.; Rao, N. S.; Tan, R.; Muhandiram, D. R. et al. Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. *Science* **1996**, *273*, 1547-1551.
(26) Puglisi, J. D.; Chen, L.; Blanchard, S.; Frankel, A. D. Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. *Science* **1995**, *270*, 1200-1203.
(27) Legault, P.; Li, J.; Mogridge, J.; Kay, L. E.; Greenblatt, J. NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell* **1998**, *93*, 289-299.
(28) Jiang, F.; Gorin, A.; Hu, W.; Majumdar, A.; Baskerville, S. et al. Anchoring an extended HTLV-1 Rex peptide within an RNA major groove containing junctional base triples. *Structure* **1999**, *7*, 1461-1472.
(29) Hermann, T. Aminoglycoside antibiotics: old drugs and new therapeutic approaches. *Cell Mol Life Sci* **2007**, *64*, 1841-1852.
(30) Dallas, A.; Vlassov, A. V. RNAi: a novel antisense technology and its therapeutic potential. *Med Sci Monit* **2006**, *12*, RA67-74.
(31) Gallego, J.; Varani, G. Targeting RNA with small-molecule drugs: therapeutic promise and chemical challenges. *Acc Chem Res* **2001**, *34*, 836-843.
(32) Sutcliffe, J. A. Improving on nature: antibiotics that target the ribosome. *Curr Opin Microbiol* **2005**, *8*, 534-542.
(33) Hamasaki, K.; Ueno, A. Aminoglycoside antibiotics, neamine and its derivatives as potent inhibitors for the RNA-protein interactions derived from HIV-1 activators. *Bioorg Med Chem Lett* **2001**, *11*, 591-594.
(34) Frankel, A. D. Fitting peptides into the RNA world. *Curr Opin Struct Biol* **2000**, *10*, 332-340.
(35) Perez, J. J.; Corcho, F.; Llorens, O. Molecular modeling in the design of peptidomimetics and peptide surrogates. *Curr Med Chem* **2002**, *9*, 2209-2229.
(36) Blount, K. F.; Tor, Y. A tale of two targets: differential RNA selectivity of nucleobase-aminoglycoside conjugates. *Chembiochem* **2006**, *7*, 1612-1621.
(37) Jin, E.; Katritch, V.; Olson, W. K.; Kharatisvili, M.; Abagyan, R. et al. Aminoglycoside binding in the major groove of duplex RNA: the thermodynamic and electrostatic forces that govern recognition. *J Mol Biol* **2000**, *298*, 95-110.
(38) Griffey, R. H.; Hofstadler, S. A.; Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T. Determinants of aminoglycoside-binding specificity for rRNA by using mass spectrometry. *Proc Natl Acad Sci U S A* **1999**, *96*, 10129-10133.

(39) Lynch, S. R.; Puglisi, J. D. Structural origins of aminoglycoside specificity for prokaryotic ribosomes. *J Mol Biol* **2001**, *306*, 1037-1058.
(40) Sucheck, S. J.; Greenberg, W. A.; Tolbert, T. J.; Wong, C. H. Design of Small Molecules That Recognize RNA: Development of Aminoglycosides as Potential Antitumor Agents That Target Oncogenic RNA Sequences This work was supported by the NIH. We thank Professor Peter Voght for his suggestion of the oncogenic RNA sequences as targets. *Angew Chem Int Ed Engl* **2000**, *39*, 1080-1084.
(41) Kirk, S. R.; Tor, Y. tRNA(Phe) binds aminoglycoside antibiotics. *Bioorg Med Chem* **1999**, *7*, 1979-1991.
(42) Walter, F.; Vicens, Q.; Westhof, E. Aminoglycoside-RNA interactions. *Curr Opin Chem Biol* **1999**, *3*, 694-704.
(43) Vicens, Q.; Westhof, E. RNA as a drug target: the case of aminoglycosides. *Chembiochem* **2003**, *4*, 1018-1023.
(44) Luedtke, N. W.; Liu, Q.; Tor, Y. RNA-ligand interactions: affinity and specificity of aminoglycoside dimers and acridine conjugates to the HIV-1 Rev response element. *Biochemistry* **2003**, *42*, 11391-11403.
(45) Ryu, D. H.; Rando, R. R. Aminoglycoside binding to human and bacterial A-Site rRNA decoding region constructs. *Bioorg Med Chem* **2001**, *9*, 2601-2608.
(46) Kaul, M.; Barbieri, C. M.; Pilch, D. S. Defining the basis for the specificity of aminoglycoside-rRNA recognition: a comparative study of drug binding to the A sites of Escherichia coli and human rRNA. *J Mol Biol* **2005**, *346*, 119-134.
(47) Krebs, A.; Ludwig, V.; Boden, O.; Gobel, M. W. Targeting the HIV trans-activation responsive region--approaches towards RNA-binding drugs. *Chembiochem* **2003**, *4*, 972-978.
(48) Zapp, M. L.; Stern, S.; Green, M. R. Small molecules that selectively block RNA binding of HIV-1 Rev protein inhibit Rev function and viral production. *Cell* **1993**, *74*, 969-978.
(49) Ryu, D. H.; Litovchick, A.; Rando, R. R. Stereospecificity of aminoglycoside-ribosomal interactions. *Biochemistry* **2002**, *41*, 10499-10509.
(50) Garbesi, A.; Hamy, F.; Maffini, M.; Albrecht, G.; Klimkait, T. TAR-RNA binding by HIV-1 Tat protein is selectively inhibited by its L-enantiomer. *Nucl Acids Res* **1998**, *26*, 2886-2890.
(51) Huq, I.; Ping, Y. H.; Tamilarasu, N.; Rana, T. M. Controlling human immunodeficiency virus type 1 gene expression by unnatural peptides. *Biochemistry* **1999**, *38*, 5172-5177.
(52) Litovchick, A.; Rando, R. R. Stereospecificity of short Rev-derived peptide interactions with RRE IIB RNA. *RNA* **2003**, *9*, 937-948.
(53) Blount, K. F.; Zhao, F.; Hermann, T.; Tor, Y. Conformational constraint as a means for understanding RNA-aminoglycoside specificity. *J Am Chem Soc* **2005**, *127*, 9818-9829.
(54) Athanassiou, Z.; Dias, R. L.; Moehle, K.; Dobson, N.; Varani, G. et al. Structural mimicry of retroviral tat proteins by constrained beta-hairpin peptidomimetics: ligands with high affinity and selectivity for viral TAR RNA regulatory elements. *J Am Chem Soc* **2004**, *126*, 6906-6913.
(55) Karn, J. Tackling Tat. *J Mol Biol* **1999**, *293*, 235-254.
(56) Stevens, M.; De Clercq, E.; Balzarini, J. The regulation of HIV-1 transcription: molecular targets for chemotherapeutic intervention. *Med Res Rev* **2006**, *26*, 595-625.
(57) Barboric, M.; Peterlin, B. M. A new paradigm in eukaryotic biology: HIV Tat and the control of transcriptional elongation. *PLoS Biol* **2005**, *3*, e76.

(58) Jones, K. A.; Peterlin, B. M. Control of RNA initiation and elongation at the HIV-1 promoter. *Annu Rev Biochem* **1994**, *63*, 717-743.
(59) Liang, C.; Wainberg, M. A. The role of Tat in HIV-1 replication: an activator and/or a suppressor? *AIDS Rev* **2002**, *4*, 41-49.
(60) Gonda, M. A.; Luther, D. G.; Fong, S. E.; Tobin, G. J. Bovine immunodeficiency virus: molecular biology and virus-host interactions. *Virus Res* **1994**, *32*, 155-181.
(61) Greenbaum, N. L. How Tat targets TAR: structure of the BIV peptide-RNA complex. *Structure* **1996**, *4*, 5-9.
(62) Ye, X.; Kumar, R. A.; Patel, D. J. Molecular recognition in the bovine immunodeficiency virus Tat peptide-TAR RNA complex. *Chem Biol* **1995**, *2*, 827-840.
(63) Calnan, B. J.; Tidor, B.; Biancalana, S.; Hudson, D.; Frankel, A. D. Arginine-mediated RNA recognition: the arginine fork. *Science* **1991**, *252*, 1167-1171.
(64) Edwards, T. E.; Robinson, B. H.; Sigurdsson, S. T. Identification of amino acids that promote specific and rigid TAR RNA-tat protein complex formation. *Chem Biol* **2005**, *12*, 329-337.
(65) Tan, R.; Frankel, A. D. Structural variety of arginine-rich RNA-binding peptides. *Proc Natl Acad Sci U S A* **1995**, *92*, 5282-5286.
(66) Moras, D.; Poterszman, A. Getting into the major groove. Protein-RNA interactions. *Curr Biol* **1996**, *6*, 530-532.
(67) Lim, A. C.; Barton, J. K. Targeting the Tat-binding site of bovine immunodeficiency virus TAR RNA with a shape-selective rhodium complex. *Bioorg Med Chem* **1997**, *5*, 1131-1136.
(68) St-Louis, M. C.; Cojocariu, M.; Archambault, D. The molecular biology of bovine immunodeficiency virus: a comparison with other lentiviruses. *Anim Health Res Rev* **2004**, *5*, 125-143.
(69) Bogerd, H. P.; Wiegand, H. L.; Bieniasz, P. D.; Cullen, B. R. Functional differences between human and bovine immunodeficiency virus Tat transcription factors. *J Virol* **2000**, *74*, 4666-4671.
(70) Puglisi, J. D.; Tan, R.; Calnan, B. J.; Frankel, A. D.; Williamson, J. R. Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science* **1992**, *257*, 76-80.
(71) Brodsky, A. S.; Williamson, J. R. Solution structure of the HIV-2 TAR-argininamide complex. *J Mol Biol* **1997**, *267*, 624-639.
(72) Aboul-ela, F.; Karn, J.; Varani, G. The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J Mol Biol* **1995**, *253*, 313-332.
(73) Long, K. S.; Crothers, D. M. Characterization of the solution conformations of unbound and Tat peptide-bound forms of HIV-1 TAR RNA. *Biochemistry* **1999**, *38*, 10059-10069.
(74) Barboric, M.; Taube, R.; Nekrep, N.; Fujinaga, K.; Peterlin, B. M. Binding of Tat to TAR and recruitment of positive transcription elongation factor b occur independently in bovine immunodeficiency virus. *J Virol* **2000**, *74*, 6039-6044.
(75) Garber, M. E.; Wei, P.; KewalRamani, V. N.; Mayall, T. P.; Herrmann, C. H. et al. The interaction between HIV-1 Tat and human cyclin T1 requires zinc and a critical cysteine residue that is not conserved in the murine CycT1 protein. *Genes Dev* **1998**, *12*, 3512-3527.

(76) Richter, S.; Cao, H.; Rana, T. M. Specific HIV-1 TAR RNA loop sequence and functional groups are required for human cyclin T1-Tat-TAR ternary complex formation. *Biochemistry* **2002**, *41*, 6391-6397.

(77) Mei, H. Y.; Mack, D. P.; Galan, A. A.; Halim, N. S.; Heldsinger, A. et al. Discovery of selective, small-molecule inhibitors of RNA complexes--I. The Tat protein/TAR RNA complexes required for HIV-1 transcription. *Bioorg Med Chem* **1997**, *5*, 1173-1184.

(78) Mei, H. Y.; Cui, M.; Heldsinger, A.; Lemrow, S. M.; Loo, J. A. et al. Inhibitors of protein-RNA complexation that target the RNA: specific recognition of human immunodeficiency virus type 1 TAR RNA by small organic molecules. *Biochemistry* **1998**, *37*, 14204-14212.

(79) Cecchetti, V.; Parolin, C.; Moro, S.; Pecere, T.; Filipponi, E. et al. 6-Aminoquinolones as new potential anti-HIV agents. *J Med Chem* **2000**, *43*, 3799-3802.

(80) Richter, S.; Parolin, C.; Gatto, B.; Del Vecchio, C.; Brocca-Cofano, E. et al. Inhibition of human immunodeficiency virus type 1 tat-trans-activation-responsive region interaction by an antiviral quinolone derivative. *Antimicrob Agents Chemother* **2004**, *48*, 1895-1899.

(81) Filikov, A. V.; Mohan, V.; Vickers, T. A.; Griffey, R. H.; Cook, P. D. et al. Identification of ligands for RNA targets via structure-based virtual screening: HIV-1 TAR. *J Comput Aided Mol Des* **2000**, *14*, 593-610.

(82) Lind, K. E.; Du, Z.; Fujinaga, K.; Peterlin, B. M.; James, T. L. Structure-based computational database screening, in vitro assay, and NMR assessment of compounds that target TAR RNA. *Chem Biol* **2002**, *9*, 185-193.

(83) Du, Z.; Lind, K. E.; James, T. L. Structure of TAR RNA complexed with a Tat-TAR interaction nanomolar inhibitor that was identified by computational screening. *Chem Biol* **2002**, *9*, 707-712.

(84) Hamy, F.; Brondani, V.; Florsheimer, A.; Stark, W.; Blommers, M. J. et al. A new class of HIV-1 Tat antagonist acting through Tat-TAR inhibition. *Biochemistry* **1998**, *37*, 5086-5095.

(85) Yuan, D.; He, M.; Pang, R.; Lin, S. S.; Li, Z. et al. The design, synthesis, and biological evaluation of novel substituted purines as HIV-1 Tat-TAR inhibitors. *Bioorg Med Chem* **2007**, *15*, 265-272.

(86) Tamilarasu, N.; Huq, I.; Rana, T. M. Targeting RNA with peptidomimetic oligomers in human cells. *Bioorg Med Chem Lett* **2001**, *11*, 505-507.

(87) Hwang, S.; Tamilarasu, N.; Ryan, K.; Huq, I.; Richter, S. et al. Inhibition of gene expression in human cells through small molecule-RNA interactions. *Proc Natl Acad Sci U S A* **1999**, *96*, 12997-13002.

(88) Tamilarasu, N.; Huq, I.; Rana, T. M. Design, synthesis, and biological activity of a cyclic peptide: an inhibitor of HIV-1 tat-TAR interactions in human cells. *Bioorg Med Chem Lett* **2000**, *10*, 971-974.

(89) Hamy, F.; Felder, E. R.; Heizmann, G.; Lazdins, J.; Aboul-ela, F. et al. An inhibitor of the Tat/TAR RNA interaction that effectively suppresses HIV-1 replication. *Proc Natl Acad Sci U S A* **1997**, *94*, 3548-3553.

(90) Daelemans, D.; Schols, D.; Witvrouw, M.; Pannecouque, C.; Hatse, S. et al. A second target for the peptoid Tat/transactivation response element inhibitor CGP64222: inhibition of human immunodeficiency virus replication by blocking CXC-chemokine receptor 4-mediated virus entry. *Mol Pharmacol* **2000**, *57*, 116-124.

(91) De Clercq, E.; Schols, D. Inhibition of HIV infection by CXCR4 and CCR5 chemokine receptor antagonists. *Antivir Chem Chemother* **2001**, *12 Suppl 1*, 19-31.

(92) Runyon, S. T.; Puglisi, J. D. Design of a cyclic peptide that targets a viral RNA. *J Am Chem Soc* **2003**, *125*, 15704-15705.

(93) Leeper, T. C.; Athanassiou, Z.; Dias, R. L.; Robinson, J. A.; Varani, G. TAR RNA recognition by a cyclic peptidomimetic of Tat protein. *Biochemistry* **2005**, *44*, 12362-12372.

(94) Jiang, L., Moehle, K., Dhanapal, B., Obrecht, D., Robinson, J.A. Combinatorial Biomimetic Chemistry: Parallel Synthesis of a Small Library of beta-Hairpin Mimetics Based on Loop III from Human Platelet-Derived Growth Factor B. *Helv Chim Acta* **2000**, *83*, 3097-3112.

(95) Revzin, A. Gel electrophoresis assays for DNA-protein interactions. *Biotechniques* **1989**, *7*, 346-355.

(96) Park, Y. W.; Wilusz, J.; Katze, M. G. Regulation of eukaryotic protein synthesis: selective influenza viral mRNA translation is mediated by the cellular RNA-binding protein GRSF-1. *Proc Natl Acad Sci U S A* **1999**, *96*, 6694-6699.

(97) Feichtinger, K. Diprotected Trityl-Guanidines: A New Class of Guanidinylation Agants. *J Org Chem* **1998**, *63*, 3804.

(98) Tamaki, M.; Han, G.; Hruby, V. J. Practical and efficient synthesis of orthogonally protected constrained 4-guanidinoprolines. *J Org Chem* **2001**, *66*, 1038-1042.

(99) Fujinaga, K.; Irwin, D.; Taube, R.; Zhang, F.; Geyer, M. et al. A minimal chimera of human cyclin T1 and tat binds TAR and activates human immunodeficiency virus transcription in murine cells. *J Virol* **2002**, *76*, 12934-12939.

(100) Guntert, P.; Mumenthaler, C.; Wuthrich, K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol* **1997**, *273*, 283-298.

(101) Koradi, R.; Billeter, M.; Wuthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* **1996**, *14*, 51-55, 29-32.

(102) Griffiths-Jones, S. R.; Maynard, A. J.; Searle, M. S. Dissecting the stability of a beta-hairpin peptide that folds in water: NMR and molecular dynamics analysis of the beta-turn and beta-strand contributions to folding. *J Mol Biol* **1999**, *292*, 1051-1069.

(103) Maynard, A. J., Sharman, G. J., and Searle, M. S., Origin of b-hairpin stability in solution: Structural and thermodynamic analysis of the folding of a model peptide supports hydrophobic stabilization in water. *J Am Chem Soc* **1998**, *120*, 1996-2007.

(104) Baxter, N. J.; Williamson, M. P. Temperature dependence of 1H chemical shifts in proteins. *J Biomol NMR* **1997**, *9*, 359-369.

(105) Englander, S. W.; Mayne, L. Protein folding studied using hydrogen-exchange labeling and two-dimensional NMR. *Annu Rev Biophys Biomol Struct* **1992**, *21*, 243-265.

(106) Aboul-ela, F.; Karn, J.; Varani, G. Structure of HIV-1 TAR RNA in the absence of ligands reveals a novel conformation of the trinucleotide bulge. *Nucl Acids Res* **1996**, *24*, 3974-3981.

(107) Nakanishi, E. K., Berova, N., Woody, R.W. *Circular Dichroism*; VCH: New York, 1994.

(108) Richard, J. P.; Melikov, K.; Vives, E.; Ramos, C.; Verbeure, B. et al. Cell-penetrating peptides. A reevaluation of the mechanism of cellular uptake. *J Biol Chem* **2003**, *278*, 585-590.
(109) Frankel, A. D.; Pabo, C. O. Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* **1988**, *55*, 1189-1193.
(110) Vives, E.; Brodin, P.; Lebleu, B. A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. *J Biol Chem* **1997**, *272*, 16010-16017.
(111) Futaki, S.; Suzuki, T.; Ohashi, W.; Yagami, T.; Tanaka, S. et al. Arginine-rich peptides. An abundant source of membrane-permeable peptides having potential as carriers for intracellular protein delivery. *J Biol Chem* **2001**, *276*, 5836-5840.
(112) Suzuki, T.; Futaki, S.; Niwa, M.; Tanaka, S.; Ueda, K. et al. Possible existence of common internalization mechanisms among arginine-rich peptides. *J Biol Chem* **2002**, *277*, 2437-2443.
(113) Futaki, S.; Goto, S.; Sugiura, Y. Membrane permeability commonly shared among arginine-rich peptides. *J Mol Recognit* **2003**, *16*, 260-264.
(114) Futaki, S. Membrane-permeable arginine-rich peptides and the translocation mechanisms. *Adv Drug Deliv Rev* **2005**, *57*, 547-558.
(115) Gupta, B.; Levchenko, T. S.; Torchilin, V. P. Intracellular delivery of large molecules and small particles by cell-penetrating proteins and peptides. *Adv Drug Deliv Rev* **2005**, *57*, 637-651.
(116) Duchardt, F.; Fotin-Mleczek, M.; Schwarz, H.; Fischer, R.; Brock, R. A comprehensive model for the cellular uptake of cationic cell-penetrating peptides. *Traffic* **2007**, *8*, 848-866.
(117) Gump, J. M.; Dowdy, S. F. TAT transduction: the molecular mechanism and therapeutic prospects. *Trends Mol Med* **2007**, *13*, 443-448.
(118) Sjoback, R.; Nygren, J.; Kubista, M. Absorption and fluorescence properties of fluorescein. *Spectrochimica Acta, Part A* **1995**, *51*, L7-L21.
(119) Bonnet, D. G., C., Rousselot-Pailley, P., Joly, P., Bourel-Bonnet, L., Santraine, V., Gras-Messe, H., Melnyk, O., Solid-phase functionalization of peptides by an alpha-hydrazinoacetyl group. *J Org Chem* **2003**, *63*, 7033-7040.
(120) Bonnet, D., Ollivier, N., Gras-Masse, H., Melnyk, O. Chemoselective acylation of fully deprotected hydrazino acetyl peptides. Application to the synthesis of lipopetides. *J Org Chem* **2001**, *66*, 443-449.
(121) Robinson, J. K., Lee V., Claridge, T.D.W., Baldwin, J.E., Schofield, Ch.J. Synthesis of (2S, 3R, 4S), (2S, 3S, 4R)-epoxyprolines and aminohydroxyprolines. *Tetrahedron* **1998**, *54*, 981-996.
(122) Hoffmann, R. V., Kim, Hwa-Ok The Preparation of 2-Hydrazinyl Esters in High Optical Purity from 2-Sulfonyloxy Esters. *Tetrahedron Letters* **1990**, *31*, 2953-2956.
(123) Lerner, C., Siegrist, R., Schweizer, E., Diederich, F., Gramlich, V., Jakob-Roetne, R., Zürcher, G., Borroni, E. Bisubstrate Inhibitors for the Enzyme Catechol O-Methyltransferase (COMT): Dramatic Effects of Ribose Modifications on Binding Affinity and Binding Mode. *Helv Chim Acta* **2003**, *86*, 1045-1062.
(124) Erlanger, B. F.; Kokowsky, N.; Cohen, W. The preparation and properties of two new chromogenic substrates of trypsin. *Arch Biochem Biophys* **1961**, *95*, 271-278.

(125) Temporini, C.; Perani, E.; Mancini, F.; Bartolini, M.; Calleri, E. et al. Optimization of a trypsin-bioreactor coupled with high-performance liquid chromatography-electrospray ionization tandem mass spectrometry for quality control of biotechnological drugs. *J Chromatogr A* **2006**, *1120*, 121-131.

(126) Wang, D.; Liao, W.; Arora, P. S. Enhanced metabolic stability and protein-binding properties of artificial alpha helices derived from a hydrogen-bond surrogate: application to Bcl-xL. *Angew Chem Int Ed Engl* **2005**, *44*, 6525-6529.

(127) DeMarco, S. J.; Henze, H.; Lederer, A.; Moehle, K.; Mukherjee, R. et al. Discovery of novel, highly potent and selective beta-hairpin mimetic CXCR4 inhibitors with excellent anti-HIV activity and pharmacokinetic profiles. *Bioorg Med Chem* **2006**, *14*, 8396-8404.

(128) Janssen, S.; Jakobsen, C. M.; Rosen, D. M.; Ricklis, R. M.; Reineke, U. et al. Screening a combinatorial peptide library to develop a human glandular kallikrein 2-activated prodrug as targeted therapy for prostate cancer. *Mol Cancer Ther* **2004**, *3*, 1439-1450.

(129) Chen, L.; Frankel, A. D. A peptide interaction in the major groove of RNA resembles protein interactions in the minor groove of DNA. *Proc Natl Acad Sci U S A* **1995**, *92*, 5077-5081.

(130) Pollard, V. W.; Malim, M. H. The HIV-1 Rev protein. *Annu Rev Microbiol* **1998**, *52*, 491-532.

(131) Bartel, D. P.; Zapp, M. L.; Green, M. R.; Szostak, J. W. HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* **1991**, *67*, 529-536.

(132) Heguy, A. Inhibition of the HIV Rev transactivator : a new target for therapeutic intervention. *Front Biosci* **1997**, *2*, d283-297.

(133) Ye, X.; Gorin, A.; Ellington, A. D.; Patel, D. J. Deep penetration of an alpha-helix into a widened RNA major groove in the HIV-1 rev peptide-RNA aptamer complex. *Nat Struct Biol* **1996**, *3*, 1026-1033.

(134) Tan, R.; Frankel, A. D. Costabilization of peptide and RNA structure in an HIV Rev peptide-RRE complex. *Biochemistry* **1994**, *33*, 14579-14585.

(135) Grate, D.; Wilson, C. Role REVersal: understanding how RRE RNA binds its peptide ligand. *Structure* **1997**, *5*, 7-11.

(136) Mann, D. A.; Mikaelian, I.; Zemmel, R. W.; Green, S. M.; Lowe, A. D. et al. A molecular rheostat. Co-operative rev binding to stem I of the rev-response element modulates human immunodeficiency virus type-1 late gene expression. *J Mol Biol* **1994**, *241*, 193-207.

(137) Cullen, B. R. Nuclear mRNA export: insights from virology. *Trends Biochem Sci* **2003**, *28*, 419-424.

(138) Cullen, B. R. HIV-1 auxiliary proteins: making connections in a dying cell. *Cell* **1998**, *93*, 685-692.

(139) Schroder, H. C., Ushijima, H., Bek, A., Merz, H., Pfeifer, K., Muller, W.E.G. Inhibition of formation of Rev-RRE complex by pyronin Y. *Antivir Chem Chemother* **1993**, *4*, 103-111.

(140) Zapp, M. L.; Young, D. W.; Kumar, A.; Singh, R.; Boykin, D. W. et al. Modulation of the Rev-RRE interaction by aromatic heterocyclic compounds. *Bioorg Med Chem* **1997**, *5*, 1149-1155.

(141) DeJong, E. S.; Chang, C. E.; Gilson, M. K.; Marino, J. P. Proflavine acts as a Rev inhibitor by targeting the high-affinity Rev binding site of the Rev responsive element of HIV-1. *Biochemistry* **2003**, *42*, 8035-8046.

(142) Wang, Y.; Hamasaki, K.; Rando, R. R. Specificity of aminoglycoside binding to RNA constructs derived from the 16S rRNA decoding region and the HIV-RRE activator region. *Biochemistry* **1997**, *36*, 768-779.
(143) Qian-Cutrone, J.; Huang, S.; Trimble, J.; Li, H.; Lin, P. F. et al. Niruriside, a new HIV REV/RRE binding inhibitor from Phyllanthus niruri. *J Nat Prod* **1996**, *59*, 196-199.
(144) Chapman, R. L.; Stanley, T. B.; Hazen, R.; Garvey, E. P. Small molecule modulators of HIV Rev/Rev response element interaction identified by random screening. *Antiviral Res* **2002**, *54*, 149-162.
(145) Chaloin, L.; Smagulova, F.; Hariton-Gazal, E.; Briant, L.; Loyter, A. et al. Potent inhibition of HIV-1 replication by backbone cyclic peptides bearing the Rev arginine rich motif. *J Biomed Sci* **2007**, *14*, 565-584.
(146) Mills, N. L.; Daugherty, M. D.; Frankel, A. D.; Guy, R. K. An alpha-helical peptidomimetic inhibitor of the HIV-1 Rev-RRE interaction. *J Am Chem Soc* **2006**, *128*, 3496-3497.
(147) Fasan, R.; Dias, R. L.; Moehle, K.; Zerbe, O.; Vrijbloed, J. W. et al. Using a beta-hairpin to mimic an alpha-helix: cyclic peptidomimetic inhibitors of the p53-HDM2 protein-protein interaction. *Angew Chem Int Ed Engl* **2004**, *43*, 2109-2112.
(148) Calnan, B. J.; Biancalana, S.; Hudson, D.; Frankel, A. D. Analysis of arginine-rich peptides from the HIV Tat protein reveals unusual features of RNA-protein recognition. *Genes Dev* **1991**, *5*, 201-210.
(149) Kjems, J.; Calnan, B. J.; Frankel, A. D.; Sharp, P. A. Specific binding of a basic peptide from HIV-1 Rev. *EMBO J* **1992**, *11*, 1119-1129.
(150) Frank, D. N.; Pace, N. R. Ribonuclease P: unity and diversity in a tRNA processing ribozyme. *Annual Review of Biochemistry* **1998**, *67*, 153-180.
(151) Torres-Larios, A.; Swinger, K. K.; Pan, T.; Mondragon, A. Structure of ribonuclease P--a universal ribozyme. *Current Opinion in Structural Biology* **2006**, *16*, 327-335.
(152) Christian, E. L.; Zahler, N. H.; Kaye, N. M.; Harris, M. E. Analysis of substrate recognition by the ribonucleoprotein endonuclease RNase P. *Methods (Duluth)* **2002**, *28*, 307-322.
(153) Evans, D.; Marquez, S. M.; Pace, N. R. RNase P: interface of the RNA and protein worlds. *Trends Biochem Sci* **2006**, *31*, 333-341.
(154) Reich, C.; Olsen, G. J.; Pace, B.; Pace, N. R. Role of the protein moiety of ribonuclease P, a ribonucleoprotein enzyme. *Science* **1988**, *239*, 178-181.
(155) Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **1983**, *35*, 849-857.
(156) Buck, A. H.; Kazantsev, A. V.; Dalby, A. B.; Pace, N. R. Structural perspective on the activation of RNAse P RNA by protein. *Nat Struct Mol Biol* **2005**, *12*, 958-964.
(157) Kikovska, E.; Svard, S. G.; Kirsebom, L. A. Eukaryotic RNase P RNA mediates cleavage in the absence of protein. *Proc Natl Acad Sci U S A* **2007**, *104*, 2062-2067.
(158) Buck, A. H.; Dalby, A. B.; Poole, A. W.; Kazantsev, A. V.; Pace, N. R. Protein activation of a ribozyme: the role of bacterial RNase P protein. *Embo J* **2005**, *24*, 3360-3368.
(159) Kazantsev, A. V.; Pace, N. R. Bacterial RNase P: a new view of an ancient enzyme. *Nat Rev Microbiol* **2006**, *4*, 729-740.

(160) Kirsebom, L. A. RNase P RNA mediated cleavage: substrate recognition and catalysis. *Biochimie* **2007**, *89*, 1183-1194.
(161) Haas, E. S.; Brown, J. W. Evolutionary variation in bacterial RNase P RNAs. *Nucl Acids Res* **1998**, *26*, 4093-4099.
(162) Gopalan, V. Uniformity amid diversity in RNase P. *Proc Natl Acad Sci U S A* **2007**, *104*, 2031-2032.
(163) Harris, J. K.; Haas, E. S.; Williams, D.; Frank, D. N.; Brown, J. W. New insight into RNase P RNA structure from comparative analysis of the archaeal RNA. *Rna* **2001**, *7*, 220-232.
(164) Krasilnikov, A. S.; Yang, X.; Pan, T.; Mondragon, A. Crystal structure of the specificity domain of ribonuclease P. *Nature* **2003**, *421*, 760-764.
(165) Krasilnikov, A. S.; Xiao, Y.; Pan, T.; Mondragon, A. Basis for structural diversity in homologous RNAs. *Science* **2004**, *306*, 104-107.
(166) Torres-Larios, A.; Swinger, K. K.; Krasilnikov, A. S.; Pan, T.; Mondragon, A. Crystal structure of the RNA component of bacterial ribonuclease P. *Nature* **2005**, *437*, 584-587.
(167) Kazantsev, A. V.; Krivenko, A. A.; Harrington, D. J.; Holbrook, S. R.; Adams, P. D. et al. Crystal structure of a bacterial ribonuclease P RNA. *Proc Natl Acad Sci U S A* **2005**, *102*, 13392-13397.
(168) Stams, T.; Niranjanakumari, S.; Fierke, C. A.; Christianson, D. W. Ribonuclease P protein structure: evolutionary origins in the translational apparatus. *Science* **1998**, *280*, 752-755.
(169) Spitzfaden, C.; Nicholson, N.; Jones, J. J.; Guth, S.; Lehr, R. et al. The structure of ribonuclease P protein from Staphylococcus aureus reveals a unique binding site for single-stranded RNA. *J Mol Biol* **2000**, *295*, 105-115.
(170) Kazantsev, A. V.; Krivenko, A. A.; Harrington, D. J.; Carter, R. J.; Holbrook, S. R. et al. High-resolution structure of RNase P protein from Thermotoga maritima. *Proc Natl Acad Sci U S A* **2003**, *100*, 7497-7502.
(171) Pannucci, J. A.; Haas, E. S.; Hall, T. A.; Harris, J. K.; Brown, J. W. RNase P RNAs from some Archaea are catalytically active. *Proc Natl Acad Sci U S A* **1999**, *96*, 7803-7808.
(172) Tsai, H. Y.; Pulukkunat, D. K.; Woznick, W. K.; Gopalan, V. Functional reconstitution and characterization of Pyrococcus furiosus RNase P. *Proc Natl Acad Sci U S A* **2006**, *103*, 16147-16152.
(173) Robertson, H. D.; Altman, S.; Smith, J. D. Purification and properties of a specific Escherichia coli ribonuclease which cleaves a tyrosine transfer ribonucleic acid presursor. *J Biol Chem* **1972**, *247*, 5243-5251.
(174) Altman, S.; Robertson, H. D. RNA precursor molecules and ribonucleases in E. coli. *Mol Cell Biochem* **1973**, *1*, 83-93.
(175) Sharkady, S. M.; Nolan, J. M. Bacterial ribonuclease P holoenzyme crosslinking analysis reveals protein interaction sites on the RNA subunit. *Nucleic Acids Res* **2001**, *29*, 3848-3856.
(176) Biswas, R.; Ledman, D. W.; Fox, R. O.; Altman, S.; Gopalan, V. Mapping RNA-protein interactions in ribonuclease P from Escherichia coli using disulfide-linked EDTA-Fe. *J Mol Biol* **2000**, *296*, 19-31.
(177) Niranjanakumari, S.; Stams, T.; Crary, S. M.; Christianson, D. W.; Fierke, C. A. Protein component of the ribozyme ribonuclease P alters substrate recognition by directly contacting precursor tRNA. *Proc Natl Acad Sci U S A* **1998**, *95*, 15212-15217.

(178) Sharkady, S. M.; Nolan, J. M. Bacterial ribonuclease P holoenzyme crosslinking analysis reveals protein interaction sites on the RNA subunit. *Nucl Acids Res* **2001**, *29*, 3848-3856.
(179) Rox, C.; Feltens, R.; Pfeiffer, T.; Hartmann, R. K. Potential contact sites between the protein and RNA subunit in the Bacillus subtilis RNase P holoenzyme. *J Mol Biol* **2002**, *315*, 551-560.
(180) Tsai, H. Y.; Masquida, B.; Biswas, R.; Westhof, E.; Gopalan, V. Molecular modeling of the three-dimensional structure of the bacterial RNase P holoenzyme. *J Mol Biol* **2003**, *325*, 661-675.
(181) Niranjanakumari, S.; Day-Storms, J. J.; Ahmed, M.; Hsieh, J.; Zahler, N. H. et al. Probing the architecture of the B. subtilis RNase P holoenzyme active site by cross-linking and affinity cleavage. *RNA* **2007**, *13*, 521-535.
(182) Kim, M.; Hyun Park, B.; Lee, Y. Effects of terminal deletions in C5 protein on promoting RNase P catalysis. *Biochem Biophys Res Commun* **2000**, *268*, 118-123.
(183) Gosringer, M.; Hartmann, R. K. Function of heterologous and truncated RNase P proteins in Bacillus subtilis. *Mol Microbiol* **2007**, *66*, 801-813.
(184) Mikkelsen, N. E.; Brannvall, M.; Virtanen, A.; Kirsebom, L. A. Inhibition of RNase P RNA cleavage by aminoglycosides. *Proc Natl Acad Sci U S A* **1999**, *96*, 6155-6160.
(185) Tekos, A.; Stathopoulos, C.; Tsambaos, D.; Drainas, D. RNase P: a promising molecular target for the development of new drugs. *Curr Med Chem* **2004**, *11*, 2979-2989.
(186) Eubank, T. D.; Biswas, R.; Jovanovic, M.; Litovchick, A.; Lapidot, A. et al. Inhibition of bacterial RNase P by aminoglycoside-arginine conjugates. *FEBS Lett* **2002**, *511*, 107-112.
(187) Hori, Y.; Rogert, M. C.; Tanaka, T.; Kikuchi, Y.; Bichenkova, E. V. et al. Porphyrins and porphines bind strongly and specifically to tRNA, precursor tRNA and to M1 RNA and inhibit the ribonuclease P ribozyme reaction. *Biochim Biophys Acta* **2005**, *1730*, 47-55.
(188) Hori, Y.; Bichenkova, E. V.; Wilton, A. N.; Tanaka, T.; Douglas, K. T. et al. Porphyrins and porphines inhibit the ribonuclease P reaction in vitro. *Nucleic Acids Res Suppl* **2002**, 111-112.
(189) Hori, Y.; Bichenkova, E. V.; Wilton, A. N.; El-Attug, M. N.; Sadat-Ebrahimi, S. et al. Synthetic inhibitors of the processing of pretransfer RNA by the ribonuclease P ribozyme: enzyme inhibitors which act by binding to substrate. *Biochemistry* **2001**, *40*, 603-608.
(190) Willkomm, D. K.; Gruegelsiepe, H.; Goudinakis, O.; Kretschmer-Kazemi Far, R.; Bald, R. et al. Evaluation of bacterial RNase P RNA as a drug target. *Chembiochem* **2003**, *4*, 1041-1048.
(191) Wade, D.; Boman, A.; Wahlin, B.; Drain, C. M.; Andreu, D. et al. All-D amino acid-containing channel-forming antibiotic peptides. *Proc Natl Acad Sci U S A* **1990**, *87*, 4761-4765.
(192) Bland, J. M.; De Lucca, A. J.; Jacks, T. J.; Vigo, C. B. All-D-cecropin B: synthesis, conformation, lipopolysaccharide binding, and antibacterial activity. *Mol Cell Biochem* **2001**, *218*, 105-111.
(193) Chen, J.; Falla, T. J.; Liu, H.; Hurst, M. A.; Fujii, C. A. et al. Development of protegrins for the treatment and prevention of oral mucositis: structure-activity relationships of synthetic protegrin analogues. *Biopolymers* **2000**, *55*, 88-98.

(194) Chen, Y.; Vasil, A. I.; Rehaume, L.; Mant, C. T.; Burns, J. L. et al. Comparison of biophysical and biologic properties of alpha-helical enantiomeric antimicrobial peptides. *Chem Biol Drug Des* **2006**, *67*, 162-173.
(195) Casteels, P.; Tempst, P. Apidaecin-type peptide antibiotics function through a non-poreforming mechanism involving stereospecificity. *Biochem Biophys Res Commun* **1994**, *199*, 339-345.
(196) Fehlbaum, P.; Bulet, P.; Chernysh, S.; Briand, J. P.; Roussel, J. P. et al. Structure-activity analysis of thanatin, a 21-residue inducible insect defense peptide with sequence homology to frog skin antimicrobial peptides. *Proc Natl Acad Sci U S A* **1996**, *93*, 1221-1225.
(197) D'Amato R, F.; Thornsberry, C.; Baker, C. N.; Kirven, L. A. Effect of calcium and magnesium ions on the susceptibility of Pseudomonas species to tetracycline, gentamicin polymyxin B, and carbenicillin. *Antimicrob Agents Chemother* **1975**, *7*, 596-600.
(198) Sugiarto, H.; Yu, P. L. Effects of cations on antimicrobial activity of ostricacins-1 and 2 on E. coli O157:H7 and S. aureus 1056MRSA. *Curr Microbiol* **2007**, *55*, 36-41.
(199) Ando, T.; Tanaka, T.; Hori, Y.; Kikuchi, Y. Regulation of bacterial RNase P ribozyme reaction by divalent cation and guide DNA. *Nucl Acids Res Suppl* **2002**, 271-272.
(200) Smith, D.; Burgin, A. B.; Haas, E. S.; Pace, N. R. Influence of metal ions on the ribonuclease P reaction. Distinguishing substrate binding from catalysis. *J Biol Chem* **1992**, *267*, 2429-2436.
(201) Perin, D. D. *Purification of Laboratory Chemicals*; Oxford, 1988.
(202) Chan W. C., W. P. D. *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*; Oxford University Press, 1999; 376.
(203) Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal Biochem* **1970**, *34*, 595-598.
(204) Vojkovsky, T. Detection of secondary amines on solid phase. *Pept Res* **1995**, *8*, 236-237.
(205) Bartels, C., Xia, T.X., Billeter, M., Güntert, P., Wüthrich, K., The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J Biomol NMR* **1995**, *6*, 1-10.
(206) Seebach, D., Abele, S., Gademann, K., Guichard, G., Hintermann, T., Jaun, B., Matthews, J.L., Schreiber, J.V., Oberer, L., Ulrich Hommel, U., Widmer, H., beta2- and beta3-Peptides with Proteinaceous Side Chains: Synthesis and solution structures of constitutional isomers, a novel helical secondary structure and the influence of solvation and hydrophobic interactions on folding. *Helv Chim Acta* **1998**, *81*, 932-982.
(207) NCCLS, N. C. f. C. L. S. *Approved standard M7-A3*; Villanova, 1993.
(208) Steinberg, D. A.; Hurst, M. A.; Fujii, C. A.; Kung, A. H.; Ho, J. F. et al. Protegrin-1: a broad-spectrum, rapidly microbicidal peptide with in vivo activity. *Antimicrob Agents Chemother* **1997**, *41*, 1738-1742.
(209) Ocampo, L., Ezquerra, J.M., Digestive protease activity in juvenile Farfantepenaeus californiensis as a function of dissolved oxygen and temperature. *Aquaculture Research* **2002**, *33*, 1073-1080.
(210) Drake, B., Patek, M., Lebl, M. A convenient preparation of monosubstituted N,N'-di(Boc)-protected Guanidines. *Synthesis* **1994**, 579-582.
(211) Patchett, A. A., Witkop, B. Studies on hydroxyproline. *J Am Chem Soc* **1957**, *79*, 185-192.

(212) Remuzon, P.; Bouzard, D.; Guiol, C.; Jacquet, J. P. Fluoronaphthyridines as antibacterial agents. 6. Synthesis and structure-activity relationships of new chiral 7-(1-, 3-, 4-, and 6-methyl-2,5-diazabicyclo[2.2.1]heptan-2-yl)naphthyridine analogues of 7-[(1R,4R)-2,5- diazabicyclo[2.2.1]heptan-2-yl]-1-(1,1-dimethylethyl)-6-fluoro-1,4-dihy dro-4-oxo-1,8-naphthyridine-3-carboxylic acid. Influence of the configuration on blood pressure in dogs. A quinolone-class effect. *J Med Chem* **1992**, *35*, 2898-2909.

(213) Scott, M. G.; Hancock, R. E. Cationic antimicrobial peptides and their multifunctional role in the immune system. *Crit Rev Immunol* **2000**, *20*, 407-431.

(214) van 't Hof, W.; Veerman, E. C.; Helmerhorst, E. J.; Amerongen, A. V. Antimicrobial peptides: properties and applicability. *Biol Chem* **2001**, *382*, 597-619.

(215) Nakamura, T.; Furunaka, H.; Miyata, T.; Tokunaga, F.; Muta, T. et al. Tachyplesin, a class of antimicrobial peptide from the hemocytes of the horseshoe crab (Tachypleus tridentatus). Isolation and chemical structure. *J Biol Chem* **1988**, *263*, 16709-16713.

(216) Lehrer, R. I.; Lichtenstein, A. K.; Ganz, T. Defensins: antimicrobial and cytotoxic peptides of mammalian cells. *Annu Rev Immunol* **1993**, *11*, 105-128.

(217) Tossi, A.; Sandri, L.; Giangaspero, A. Amphipathic, alpha-helical antimicrobial peptides. *Biopolymers* **2000**, *55*, 4-30.

(218) Prenner, E. J.; Lewis, R. N.; McElhaney, R. N. The interaction of the antimicrobial peptide gramicidin S with lipid bilayer model and biological membranes. *Biochim Biophys Acta* **1999**, *1462*, 201-221.

(219) Fahrner, R. L.; Dieckmann, T.; Harwig, S. S.; Lehrer, R. I.; Eisenberg, D. et al. Solution structure of protegrin-1, a broad-spectrum antimicrobial peptide from porcine leukocytes. *Chem Biol* **1996**, *3*, 543-550.

(220) Kawano, K.; Yoneya, T.; Miyata, T.; Yoshikawa, K.; Tokunaga, F. et al. Antimicrobial peptide, tachyplesin I, isolated from hemocytes of the horseshoe crab (Tachypleus tridentatus). NMR determination of the beta-sheet structure. *J Biol Chem* **1990**, *265*, 15365-15367.

(221) Tamamura, H.; Ikoma, R.; Niwa, M.; Funakoshi, S.; Murakami, T. et al. Antimicrobial activity and conformation of tachyplesin I and its analogs. *Chem Pharm Bull (Tokyo)* **1993**, *41*, 978-980.

(222) Kokryakov, V. N.; Harwig, S. S.; Panyutich, E. A.; Shevchenko, A. A.; Aleshina, G. M. et al. Protegrins: leukocyte antimicrobial peptides that combine features of corticostatic defensins and tachyplesins. *FEBS Lett* **1993**, *327*, 231-236.

(223) Yasin, B.; Harwig, S. S.; Lehrer, R. I.; Wagar, E. A. Susceptibility of Chlamydia trachomatis to protegrins and defensins. *Infect Immun* **1996**, *64*, 709-713.

(224) Srinivas, N. New Antimicrobial beta-Hairpin Peptidomimetics, PhD Thesis. In *Mathematisch - naturwissenschaftlichen Fakultät*; Universität Zürich, 2005.

(225) Shankaramma, S. C.; Athanassiou, Z.; Zerbe, O.; Moehle, K.; Mouton, C. et al. Macrocyclic hairpin mimetics of the cationic antimicrobial peptide protegrin I: a new family of broad-spectrum antibiotics. *Chembiochem* **2002**, *3*, 1126-1133.

# Zusammenfassung

Diese Arbeit befasst sich mit der Entwicklung neuer antiviraler und antimikrobieller β-Haarnadel-Peptidmimetika (β-Hairpin). Das Ziel dabei war die Inhibition von wichtigen viralen Protein-RNA-Wechselwirkungen zwischen den Regulator-Proteinen Tat und Rev und ihren Ziel-RNAs TAR und RRE, und zwischen der RNA- und Protein-Domäne der bakteriellen RNase P. RNA und RNA/Protein-Wechselwirkungen haben in den letzten Jahrzehnten immer mehr an Bedeutung gewonnen. Die steigende Resistenz gegen antivirale Medikamente macht diese neuen Ziele interessant. Im Allgemeinen wird angenommen, dass es für Pathogene schwierig ist, RNA mit wichtigen Funktionen zu mutieren. Es wurde hier versucht die RNA-bindenden Domänen der oben genannten Proteine durch zyklische Peptide nachzuahmen, welche ein heterochirales Diprolin-Templat enthalten. Diese Peptide bilden eine interessante Klasse von Peptidmimetika, da das D-Pro-L-Pro Motiv zur Bildung von β-Hairpins führt. Diese Mimetika können aus kommerziellen Aminosäuren über standardisierte Festphasensynthese oder in Lösung hergestellt werden.

Die Wechselwirkung des viralen Tat-Proteins und seiner TAR-RNA (transactivator response element) spielt eine entscheidende Rolle im viralen Lebenszyklus. Die argininreiche RNA-bindende Domäne vom Rinder Immundefizienzvirus (BIV) Tat-Protein nimmt nach der Bindung an TAR-RNA eine β-Hairpin-Konformation an. Bisherige Studien von BIV Tat-Peptidmimetika haben die Identifikation eines sehr potenten BIV Tat-TAR Inhibitor, genannt BIV2, ermöglicht. Seine stabile β-Hairpin-Konformation und seine hohe Affinität für BIV TAR-RNA haben die Strukturaufklärung des Komplexes mit BIV TAR durch NMR-Analysen ermöglicht. In dieser Arbeit wurden ausgedehnte Struktur-Aktivitäts-Beziehungs-Studien (structure-activity-relationship, SAR) mit diesen 12-mer Loop BIV Tat-Peptidmimetika durchgeführt, welche weitere sehr potente und selektive Liganden lieferten. Weiterhin diente die BIV Tat-TAR Wechselwirkung als ein Modell, aus welchem wichtige Informationen über RNA-Erkennung gewonnen und auf den humanen Immundefizienzvirus Typ-1 (HIV-1) Tat-TAR übertragen werden konnten, was das Design neuer HIV Tat-Inhibitoren erleichterte. Durch Studien an einer umfangreichen Bibliothek aus BIV Tat-Peptidmimetika konnten einige potente

und selektive Liganden für HIV-1 TAR-RNA identifiziert werden. Der beste Ligand, **L2-15**, weist eine Dissoziationskonstante ($K_d$) von ca. 75 nM und einen $IC_{50}$ von ca. 100 nM für die Hemmung der ADP-1 Binding an TAR auf.

Analog zur BIV Anordnung, bindet das HIV Tat-Protein an die Wölbung (bulge) der TAR RNA Stamm-Schleife Struktur. Um die Funktion zu gewährleisten, muss ein ternärer Komplex zwischen Tat-TAR und Cyclin T1 gebildet werden, in welchem Cyclin T1 mit Tat und der Spitze des TAR-RNA-Loops wechselwirkt. Aufgrund von Computermodelling und SAR-Studien, wurden neue 16-mer Loop HIV-1 Tat-Peptidmimetika entwickelt, welche eine stabile β-Hairpin-Konformation in Wasser besitzen, mit hoher Affinität an TAR-RNA binden und fähig sind, die Tat-TAR Wechselwirkung zu hemmen. Das beste Peptid in dieser Klasse, **L5-11**, weist einen $K_d$-Wert von ca. 25 nM in Gegenwart eines grossen Überschusses an tRNA und ca. 0.5 nM in Abwesenheit der Konkurrenz-tRNA auf. Weil die Struktur des HIV-1 Tat-TAR-Komplexes noch nicht bekannt ist, wurden ausgedehnte SAR-Studien mit den 16-mer Loop Mimetika durchgeführt, um das Verständnis der Erkennungprinzipen dieser Protein/RNA-Wechselwirkung zu vertiefen. Weitere Studien mit HIV Tat-Peptidmimetika haben ihre Fähigkeit die Zellmembranen zu passieren und ihre Stabilität in biologischer Umgebung aufgezeigt.

Der zweite kritische Schritt in der viralen Replikation ist die Bindung des Rev-Proteins an RRE-RNA (Rev Response Element), welche den Transport ungespleisster und partiell gespleisster viraler RNA aus dem Zellkern in das Zytoplasma reguliert. Die RNA-bindende Domäne des Rev-Proteins nimmt im Komplex mit RRE eine α-helikale Konformation an. Im zweiten Teil dieser Arbeit wurde ein neues Konzept angewandt, um diese α-helikale Struktur in β-Hairpin-Peptidmimetika nachzuahmen. Diese Peptidmimetika haben ein robustes Strukturgerüst, das die für die RRE-Erkennung wichtigen Gruppen präsentiert. Durch Studien einer kleinen Familie zyklischer β-Hairpin-Peptidmimetika basierend auf BIV Tat-Peptidmimetika, und durch das Design einer neuen Serie von Peptiden basierend auf der Sequenz der RNA-bindenden Domäne des Rev-Proteins, wurde es ermöglicht, ein Set potenter RRE-Liganden zu entwickeln, welche vergleichbare $K_d$-Werte zu denen des Rev-Peptids aufzeigten. Die β-Hairpin-Konformation der Peptide wurde durch den Einbau einer Disulfidbrücke stabilisiert, was zu einem sehr aktiven Rev-Peptidmimetikum, **R2-14**, mit stabiler β-Hairpin-Struktur in wässriger Lösung führte.

Ein ähnlicher Ansatz wurde auch für das Design neuer Peptidinhibitoren der *E. coli* RNase P, einem Ribonucleoprotein verwendet, welches die Reifung der Precursor-tRNA in allen drei Lebensdomänen reguliert. Ausgedehnte biochemische Untersuchungen an zwei verwandten Bakterienstämmen, *E. coli* und *B. subtilis*, ergaben, dass α-helikale, hoch konservierte und argininreiche Domänen des Proteinteils von RNase P mit dem katalytisch aktiven RNA-Teil wechselwirken. Durch Testen der Rev-Peptidmimetika, welche Sequenzähnlichkeit mit der RNA-bindenden Domäne von RNase P haben, wurden zwei potente, gegen *E. coli* selektive, antimikrobielle Wirktstoffe gefunden. Die Informationen dieser besten Peptidmimetika und die Entwicklung eines Computermodells eines β-Hairpins, welches mit dem helikalen Epitop überlagert, führten zur Entwickelung sehr aktiver Inhibitoren von *E. coli* RNAse P. Der beste Inhibitor, **P2-05**, weist einen MIC-Wert (minimal inhibitory concentration) von 0.06 μg/ml auf. Diese Klasse von Peptiden hat günstige pharmakologische *in vitro*-Eigenschaften. Dazu gehören eine geringe haemolitische Aktivität und hohe Plasmastabilität. Diese Eigenschaften machen solche Peptidmimetika als zukünftige antimikrobielle Wirkstoffe interessant.

Die drei beschriebenen Konzepte zur Herstellung von Inhibitoren von Protein-RNA Wechselwirkungen führten zur Entdeckung sehr potenter und selektiver antiviraler und antimikrobieller Wirkstoffe, die sehr interessante Startpunkte für die Entwicklung von Medikamenten sind. Zusätzlich zu ihrem pharmakologischen Potenzial brachten ausgedehnte SAR-Studien wichtige Informationen über die Chemie und die Prinzipen der RNA-Erkennung durch diese strukturell und konformationell eingeschränkten Peptidmimetika. Diese könnten bei der Optimierung von anderen bekannten RNA-Liganden und in Zukunft bei der Entwicklung von Wirkstoffen gegen andere funktionelle RNAs helfen.

## Acknowledgments

First of all, I would like to thank my supervisor *Prof. John A. Robinson* for giving me the opportunity to work on the exciting topics which are described in this thesis. I am grateful for his guidance, support, encouragement and patience over the past four years.

I would also like to thank *Prof. Gabriele Varani* for a fruitful cooperation on BIV/HIV Tat and HIV Rev peptidomimetics project. Many thanks go also to Dr. *Zafiria Athanassiou* and *Amy Davidson* for performing EMSAs reported in this thesis.

I would like to thank *Prof. Nathan W. Luedtke* for co-refereeing this manuscript.

Special thanks go to *Dr. Kerstin Moehle* for performing the structural studies on the peptides, for her help and patience in explaining details of structural calculations.

I also wish to thank the following people for their help to complete this work:

- *Fr. Annelies Meier* for doing the antimicrobial assays reported in this thesis,
- *Dr. Annabelle Freund-Renard* for a constant support and guidance concerning chemistry and technical matters,
- *Dr. Urs Ziegler* for his introduction to the confocal microscopy technique and his help in performing the cell permeability studies,
- *Dr. Ricardo Dias* for his help, encouragement and introduction to BIV Tat project, as well as to parallel synthesis of peptide libraries,
- *Dr. Khaled Abou Hadeed* and *Prof. Stefan Bienz* for their support and excellent supervision during praktikum,
- *Arin Ghasparian* for correcting my German translation of the summary.

I wish also to thank all the members of the Robinson group (past and present) for their help and support all through my stay here. Many thanks also go to *Fr. Heike Götzmann*, *Fr. Marie-Therèse Bohley, Fr. Christa Werner* and *Fr. Cornelia Moor* for their friendliness and help in administrative matters. Many thanks also go to the NMR and MS teams at OCI, University of Zurich.

Finally I would like to thank my parents and my husband for their constant encouragement and moral support all through my academic career.

# Curriculum Vitae

Surname: PATORA-KOMISARSKA
First Name: KRYSTYNA
Date of Birth: 22.02.1978
Place of Birth: POLAND, ŁÓDŹ
Nationality: POLISH

**Education:**

**2004-2008** PhD Studies at the University of Zurich, Institute of Organic Chemistry, under supervision of Prof. J. A. Robinson; Thesis title: *Novel β-Hairpin Peptidomimetic Inhibitors of Protein-RNA Interactions in BIV/HIV-1 and in Bacterial RNase P*.

**2002** Diploma work at the Technical University of Lodz under supervision of Prof J. Zabrocki; Diploma title: *Synthesis of oxitocine and vasopresine modified by 1,5-substituted tetrazole ring*.

**2001** Socrates-Erasmus scholarship, Gent University, Belgium, Department of Organic Chemistry. Supervision of Prof. Ch. Stevens.

**1997-2003** Undergraduate Studies at Technical University of Lodz, Poland, Institute of Organic Chemistry.

**1993-1997** XXVI High School in Lodz, Poland, Matura with honours.

**Working experience:**

**2005-2007** Teaching Assistant at the University of Zurich in the practical classes of general chemistry for chemistry students and medicine students.

**2003-2004** Post-graduate position, Novartis Pharma AG Basel, Switzerland, Department of Arthritis and Bone metabolism. Supervision of Dr. P. Janser.

**Publications:**

1) Structure-Guided Peptidomimetic Design Leads to Nanomolar β-Hairpin Inhibitors of the Tat-TAR Interaction of Bovine Immunodeficiency Virus.
   Athanassiou, Z., Patora, K., Dias, R., Moehle, K., Robinson J.A., Varani, G., *Biochemistry* 2007, *46*, 741
2) Design of β-Hairpin Peptidomimetics That Inhibit Binding of α-Helical HIV-1 Rev Protein to the Rev Response Element RNA.
   Moehle, K., Athanassiou, Z., Patora, K., Davidson, A., Varani, G., Robinson, J.A., *Angew. Chem. Int. Ed.* 2007, *46*, 9101

3) Synthesis of Substituted cyclopropylphosphonates by Michael Induced Ring Closure (MIRC) Reaction.
   Stevens, Ch., Van Heecke, G., Barbero, C., Patora, K., De Kimpe, N., Verhe, R., *Synlett* 2002, *7*, 1089

**Posters:**

1) The design, synthesis and binding studies of a novel class of β-hairpin Rev peptidomimetics.
   Patora, K., Athanassiou, Z., Moehle, K., Varani, G., Robinson, J.A.,
   29$^{th}$ European Peptide Symposium, Gdansk, Poland, 2006

2) The structure-guided design and studies of β-hairpin Tat peptidomimetics.
   Patora, K., Athanassiou, Z., Moehle, K., Dias, R.L.A., Varani, G., Robinson, J.A.,
   2$^{nd}$ European Conference on Chemistry for Life Sciences, Wroclaw, Poland, 2007